A SYSTEM IDENTIFICATION APPROACH TO PROCESS-BASED PLANT GROWTH MODEL REDUCED-ORDER PARAMETER ESTIMATION

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Craig Randall Elevitch

December 2018

A SYSTEM IDENTIFICATION APPROACH TO PROCESS-BASED TREE GROWTH MODEL REDUCED-ORDER PARAMETER ESTIMATION

Craig Randall Elevitch, Ph.D.

Cornell University 2018

In order to describe underlying biophysical mechanisms, process-based plant growth models often contain an excessive number of parameters when only considering the accuracy of the model outputs. Regarding their influence on the model output, parameters of complex nonlinear plant growth models interact in ways that cannot be easily predicted based upon their roles in the component submodels describing biophysical processes. Parameter estimation in plant growth models is often challenged by lack of a means to interpret the relative importance of parameters in complex nonlinear models. In multi-crop models such as for agroforestry, increased model complexity and lack of data for novel crop combinations in varying environments further exacerbate the difficulties of discerning which parameters are important for estimation. The approach is based upon foundational system identification theory applied to a class of deterministic process-based predictive growth models. Evaluating the Hessian of the quadratic cost function determines the relative importance of parameters to its curvature. Subject to a list of model requirements, the Hessian can be computed given the input-output data and an estimated location in the parameter space provided by research into underlying biophysical processes and expert knowledge. For this analysis, field data are not required, rather, reliable simulated climate data are used to drive the model, which itself provides *a priori* output data for analysis.

The analysis method is presented as a procedure for determining a ranking of parameter importance that can be used by model developers to provide end users

with guidance for parameter estimation given real data for novel crops and crop combinations. The goal of this procedure is to arrive at a reduced-order parameter space that can be estimated entirely from input-output data. Furthermore, the goal is for the reduced model to closely follow the outputs of the original system (with any feasible parameterization) when driven by any input in the input class. The procedure is demonstrated on the well-known Yield-SAFE predictive agroforestry growth model. The advantages of an input-output system identification approach may also carry over into field trial design or model structure revisions. Further, because model parameterization can be based only on readily accessible model outputs, relatively low-tech data collection strategies emphasizing on-farm participatory research become possible. Participatory approaches allow a broader range of useful data to be collected for evaluating complex crop combinations.

Keywords: tree growth modeling; agroforestry; system identification; complexity reduction; process-based models

BIOGRAPHICAL SKETCH

Craig R. Elevitch received a BSc degree in System Science and a BA degree in Philosophy (cum Laude) from the University of California, San Diego, Revelle College in 1983. During the period 1982–1983 he attended Lund Institute of Technology, Sweden as an exchange student studying system identification and control systems, after which he continued at the Department of Automatic Control as a teaching assistant for 2 years. In 1986 he began his PhD studies at Cornell University Department of Computer and Electrical Engineering and received the MS degree in January 1988 for work on an adaptive LMS filter used in communications, after which he took a leave of absence from his engineering studies to pursue a career in agroforestry.

Craig has been an educator in agroforestry since 1991. He directs Agroforestry Net, a nonprofit educational organization dedicated to empowering people in agroforestry and ecological resource management. Craig's internationally recognized publications and workshops have guided thousands in becoming more proficient in agroforestry, particularly systems inspired by the agroforestry landscapes of the Pacific Islands. His publications include *Agroforestry Guides for Pacific Islands* (2000), *Traditional Trees of Pacific Islands: Their Culture, Environment, and Use* (2006), *Specialty Crops for Pacific Islands* (2011), and *Agroforestry Landscapes for Pacific Islands: Creating abundant and resilient food systems* (2015), all of which promote diverse agricultural systems that are environmentally and ecologically regenerative. His agroforestry publications have been downloaded millions of times. Craig has presented over 150 workshops throughout the Pacific, with over 7,000 producers and resource professionals participating since 1993.

With the backing of his Cornell advisor from the late 1980's, Dr. C. Richard Johnson, Jr., Craig returned from his Cornell leave of absence in 2014 to complete his PhD with a project that investigates system identification approaches to predictive agroforestry growth modelling.

To Ngaire Gilmour,
whose extraordinarily patient and loving support
made completion of this work possible.

ACKNOWLEDGEMENTS

My deepest and humblest thanks go to my advisor and committee chairperson Dr. C. Richard 'Rick' Johnson, Jr., who I met in 1984. Rick's unbounded support for collaborative research and innovative application of engineering principles deeply influenced my career for the past 34 years. The past 4 years of working with Rick as my PhD committee chairperson has been a master class in research methodology that has taken all areas of my professional career to higher levels.

I thank my other PhD committee members Eiljan Bitar, Adam Bojanczyk, and Jim Lassoie for their advice throughout the past 4 years, and their patience with my progress pursuing this research independently while also carrying on a full career in agroforestry research and education. Bill Sethares, who I worked with as a Cornell PhD student in 1986–87, was an extremely generous tutor in getting me up to speed on modeling approaches and simulation techniques.

I am grateful for conversations and feedback from Louise Buck, Ken Mudge, Jon Conrad, Tim Fahey, and Alan Lakso in the initial stages of my research. Meine von Noordvijk and Jim Roshetko gave me early advice regarding model complexity and predictive capabilities. Paul Burgess took the time to sit down with me to explain the philosophy behind the Yield-SAFE model, as well as provided follow-up information, including the Excel implementation that turned out to be an invaluable reference implementation. I am very thankful for brief, but very valuable exchanges with Christian Dupraz, Karel Keesman, and João Palma regarding their approaches to modeling with Yield-SAFE. Later conversations with Benoît Bayol of CentraleSupélec were helpful in providing a broader context.

The administrative staff at the Department of Electrical and Computer Engineering was exceptionally helpful. In particular, my warm thanks go to Scott Coldren and Daniel Richter. I owe a great debt of gratitude to Cornell University Graduate School for accepting my readmission application after a leave of absence of 26 years and for supporting my *in absentia* student applications, enabling me to continue my agroforestry career in the Pacific Islands during the course of this research.

My deep appreciation goes to Björn Wittenmark, Gustaf Olsson, and Karl-Johan Åström, with whose support I was able to study at Lund Institute of Technology in the early 80's. During that time, I learned from graduate students Ulf Holmberg, Mats Lilja, Per Persson, and Sven Olof Andersson Hederoth, who were always generous colleagues. At Cornell in the late 80's, Bill Sethares and Gonzalo Rey were brilliant collaborators, whose working styles influenced me for life.

In restarting my PhD research in 2014, I received valuable advice and encouragement from Roger Leakey, P.K.R. Nair, Papalii Failautusi Avegalio, Kahu Kalani Souza, and Diane Ragone. Their encouragement gave me the courage to embark on this challenging journey.

On a personal note, I acknowledge colleagues and friends who have supported the completion of my PhD research and my professional career in agroforestry over the past nearly three decades: Hector Valenzuela, Ted Radovich, Neil Logan, Sophia Bowart, JB Friday, Katie Friday, Bart Lawrence, Randy Thaman, Kalani Souza, Julie Stowell, the Kimball ʻohana, Failautusi Avegalio, the Hewahewa ʻohana, Nat Tuivavalagi, Aleta San Nicolas, Pedro and Jane Tama, Aunty Shirley Kauhaihao, Sarah Purgus, Alex Meffert, Kaleo Ten, Niki Mazaroli, Kiernan Brtalik, Peneku Kihoi, and Rockey Manning.

The idea to finish my PhD studies after a 25-year leave of absence was suggested by Betty Johnson (Rick Johnson's wife) during a casual conversation about life goals in 2013. From the first time I met Betty in the late 1980's, I remember well her supportive role in embracing all of Rick's students as family. For kicking off the completion of my PhD and for being a supportive friend for over 30 years, I extend my deepest thanks to Betty.

My partner Ngaire Gilmour stood by faithfully through the highs and lows inherent in the process of taking on a challenging project. Ngaire provided much needed warm company and comfort throughout this work, and especially, someone to listen to my eccentric musings. It would be impossible for me to sufficiently acknowledge all that Ngaire did to support me.

Finally, I acknowledge my family who instilled in me the value of learning as a life-long endeavor.

CONTENTS

1 INTRODUCTION

Tree growth models are increasingly numerous (Pretzsch et al. 2015) while also trending toward increasing complexity (Weiskittel et al. 2011, p. 228). When modeling intraspecific interactions between trees or trees and crops grown in close proximity (e.g., in tree polycultures and agroforestry), complexity can increase significantly beyond individual or monospecific tree growth models due to the interactions between species (Monteith et al. 1991, van der Werf et al. 2007). In tree modeling as well as in biophysical modeling of natural processes in general, model complexity can become problematic, particularly when accurate prediction of outcomes is of primary importance (Ljung 1999, Cox et al. 2006, Sivakumar 2008). Problems include propagation of model error between interconnected model components, parameter uncertainty, and inadequate data for parameterization and validation (Beck 1983, Young and Ratto 2009). These issues suggest that when accurate prediction (such as tree biomass) is of primary importance to the modeler, reduction of estimated parameter space is desirable (Sjöberg et al. 1995).

There are various motivations behind development of tree growth models, including interpretation of experimental results, investigation of underlying biophysical processes, and prediction of outcomes (Whisler et al. 1986). Of these three areas, the latter motivation is well suited for examination within a system identification framework and best delineates the sphere within which this work is relevant. Such systems analysis has been done for control systems since the 1970's (Balakrishnan and Peterka 1969, Åström and Eykhoff 1971, Eykhoff 1981, Ljung 1996). Malézieux et al. (2009) state regarding tree growth modeling, "There

is a need for dynamic modeling tools to evaluate how wide ranges of soil conditions, various weather sequences and different management schemes modify the yield and environmental impact of multispecies systems." Accurate prediction of tree growth in agroforestry configurations has practical importance in developing and evaluating crop combinations, management planning, and risk assessment for agroecosystems with wide-ranging potential benefits in mitigating risks associated with climate change.

This work introduces a system identification approach to reducing the space for parameter optimization for deterministic predictive tree growth models. A procedure for locating a reduced-order parameter space of such models is described, with preliminary investigative work carried out on the Yield-SAFE agroforestry model (Dupraz et al. 2005, van der Werf et al. 2007) as a test case. This approach is experimental at this point, however, examples presented here suggest that a system identification approach has potential to influence tree growth modelers' perspectives on model complexity, if not their modeling strategies. The example presented for Yield-SAFE, even though already considered a "very parameter sparse" process-based model (van der Werf et al. 2007), shows that further analysis is required for parameter identifiability in an input-output (I/O) framework. Lessons learned may extend to process-based model formulation, parameter estimation, model validation, experimental design, and overarching issues of model credibility.

1.1 Purpose of models

In general terms, a model is a system representation that allows investigation without having to run experiments on the actual system (Ljung and Glad 1994a). There are numerous situations in which conducting experiments on the real-life

system may be impractical. In predicting tree growth, for example, these challenges include the time required for observations over the time scale of tree growth, i.e., years, decades, or centuries. In order to estimate outcomes or behavior of novel tree-tree or tree-crop combinations in various configurations and environments, one would wish to run numerous experiments for various planting configurations in different environments, which becomes prohibitively expensive. Furthermore, environmental conditions may change over time (such as with climate change), meaning that the future growing environments we would like to investigate may not exist for real-life experiments today. Running experiments in the abstract through a model is the most feasible or only available opportunity in most cases. In agroforestry, this realization leads many to conclude, "Modelling is the only way to go with multispecies systems" (Malezieux et al. 2009). Echoes of this assertion can be heard throughout the literature including, "Plant growth models may be the only way to integrate over the many processes that plant physiologists may study in isolation in their laboratories" (Boote et al. 1996).

The difficulty of collecting data for natural systems such as tree growth does not mean that observational data is not needed in the modeling process. Observational data is always needed for model calibration and validation even if the model itself is developed before data is collected, as is often the case in modeling complex natural systems with long temporal scales. As will be presented below, given prior system knowledge, a range of analyses can be performed on prospective models before embarking on expensive and long duration field experiments.

Since models are a representation of a system and not the system itself, they are always imperfect. As Ljung (2010) says regarding the search for the 'best' model, "We should thus not strive for the truth, but for reasonable approximations." Models can have different forms. For example, people operate from mental models of the physical environment—opening a door, picking up an object, or catching a ball. When our mental models are accurate, we can reliably predict and control outcomes, whereas when just learning a new skill, we need to gain experience in order to build a new mental model. Via direct experience and training, indigenous peoples through the ages have developed a vast body of traditional ecological knowledge collected over generations with which to predict the course of natural processes such as tree growth (Gadgil et al. 1993, Berkes et al. 2000). One could argue that the best models available today for complex multispecies, multistrata tree plantings are the mental models of indigenous people who could be considered expert practitioners (Isaac et al. 2008).

Mathematical models of relationships between observable system quantities provide the potential for vast analytical capacity for tree growth at different levels, from molecular to broad ecosystem. There is wide consensus in crop modeling that there are no universal tree growth models (Boote et al. 1996, Malézieux et al. 2009, Weiskittel et al. 2011, Affholder et al. 2012). Each model is tailored to the specific purpose of the modeler (Table 1.1). As Whistler et al. (1986) state, "There are probably as many answers to the question 'Why build crop models?' as there are modelers." This has led to a wide array of model approaches, including in "temporal resolution (daily vs. annual vs. decadal), spatial scale (stand vs. individual tree), reliance on data (statistical vs. mechanistic), representation of competitive processes (distance-independent vs.

distance-dependent), and degree of stochasticity" (Weiskittel 2014). Additionally, thoughtfully constructed tree growth models can be valuable heuristic tools in teaching and research (Sinclair and Seligman 1996).

Every model is conceptualized for a certain set of objectives based on its own assumptions, formulations, and implementation. Mathematical models capture the system dynamics of interest, i.e., they are developed for a particular purpose such as predicting outcomes or understanding underlying processes. Stage (2003) presents a comprehensive case that, "Choice of a model ... depends on the decision space defined by the actions, indicators, ecosystem scope, and cybernetic context of the decisions."

Table 1.1. What's important to the tree growth modeler?

Model purpose		**Priority model features**	**Experimentation and validation**
Control, management, economic analysis, planning	→	Accurate prediction of outcomes	Long-term field experiments
Understanding underlying processes	→	Accurate submodels, valid connections between submodels	Controlled experiments, often in laboratory conditions
Heuristic tool, teaching, interpretation	→	Specific objectives, simplicity, transparency, robust submodels (Sinclair and Seligman 1996, p. 702)	Controlled and long-term experiments

1.2 Summary of contributions of this work

The present investigation primarily focuses on process-based models developed for the purpose of accurately predicting tree growth at the level of biomass, height, diameter at breast height (DBH), and so on. Such models are intended to be utilized in optimizing planting configuration, crop combinations, total production, and management interventions, although they may not be suitable for

investigation of underlying biophysical processes. There is a tendency to develop crop models that are constructed from submodels describing the underlying mechanisms of growth processes (e.g., Linder 1982). An unintended consequence of this is often that some of the model processes are either not observable or observable only at great expense and/or with poor accuracy. The lack of high quality observational data can lead to problems with parameter identifiability. Additionally, the model structure (connections between submodels) can inadvertently render certain parameters unidentifiable (Hengl et al. 2007). Solutions to these parameter identifiability problems include designing means to measure additional data and model reduction to a form appropriate to the information content of the data collected (Raue et al. 2011).

This research takes the latter model reduction approach by starting with a trusted, process-based growth model and data for model outputs that are easily (and hopefully cheaply) collected. This analysis systematized as a procedure determines which parameters in the model can be fixed to nominal, expert-determined values, and which can be reliably identified from the input-output data. In other words, the reduction is in the order of the parameter space to be optimized. The procedure prioritizes (and allows the user to pick for identification) the parameters that will have the most impact on improving output prediction. The result is a model that preserves its mechanistic interpretation and for which the modeler can be assured of a close output match to the full model, given adequate data (Bashir et al. 2008). Since the model is assumed to be reliable and therefore a "surrogate" for the real system (Young 2012), a close match to its outputs by the reduced model is considered a successful outcome. As a benefit, a range of parameter identification analyses can

be performed on prospective models before embarking on expensive and long duration field experiments that can inform data collection strategies and experimental design.

The concepts employed here are not new. They draw from foundational research in system identification from the 1970's and 1980's, where accurate prediction of system outputs is crucial for control and optimization purposes (Åström and Eykhoff 1971, Ljung 1996). This research borrows from the basic concepts of system identification and uses off-the-shelf, generic modeling and optimization tools in MATLAB to demonstrate the feasibility of model reduction for the plant growth modelling community, which has yet to embrace this philosophy. By using foundational knowledge from when system identification broke through in control theory and presenting the concepts in a readily accessible way, interesting advantages of a system identification perspective can be demonstrated.

Even the means of parameter space reduction used here is not new and not the most sophisticated compared with the plethora of methods that abound in the literature (Li 1992, Young 2012). Instead, the method described here is based on simple, common sense approaches and straightforward mathematics that are easily accessible to a wide audience. Reaching a wide audience in the predictive plant growth modeling community is a priority—refinement and more sophisticated approaches can be developed as the need for such becomes recognized. There are no known similar transparent, algorithmic approaches towards a system identification approach for reduced-order parameter estimation for this particular scientific community.

The model selected for demonstration in this dissertation, Yield-SAFE (van der Werf et al. 2006), may be the most widely used multistory agroforestry model

designed for predictive purposes. Yield-SAFE has been developed over the past approximately 15 years, and still lacks a systematic method to reduce the number of parameters estimated, even though it is recognized that only a subset of parameters can reasonably be estimated (Palma 2017). As the use of Yield-SAFE is expanded into new crop combinations and environments, this research is particularly timely and would not necessarily have been of interest previously.

For agroforestry and other crop growth modeling in general, outcomes prediction has become particularly timely. Agroforestry is increasingly seen as an important means to regenerate degraded lands and sequester carbon while producing economically viable crops (Verchot et al. 2007, Jose 2009, Jat et al. 2016, Hillbrand et al. 2017, Toensmeier 2016, Elevitch et al. 2018). Predictive modeling can extend existing knowledge into novel crop combinations in new environments and justify investment in agroforestry for both producers and policymakers (Malezieux et al. 2009). The urgency of reliable predictive modeling for agroforestry is amplified by recent conclusions about climate change (e.g., Schoeneberger et al. 2017, IPCC in press).

Finally, a system identification approach is also timely considering today's emphasis on participatory research combined with advances in technology. The requirement that all data are easily measurable in an input-output-based approach puts data collection within reach of farmers. With today's inexpensive measurement devices, mobile phone technologies, and internet connectivity, it is conceivable that low cost systems for data collection sufficient for complex model parameterization are within reach. This opens doors to a participatory approach, allowing farmers to design their own systems and collect data in collaboration

with researchers thereby making much more data available for model refinement. This possibility was out of reach even a few years ago.

1.3 Tree growth modeling approaches

This section presents these three modeling approaches that have been used for tree growth as context for introducing a system identification perspective

- Empirical models that generate growth curves from observed size and environmental data.
- Process-based models that are built from biophysical descriptions of underlying growth processes.
- Hybrid models combining aspects of empirical and process-based models.

From a systems modeling context, the first category is a 'black box' model, where the model structure and parameters are determined from observed measurements. The second category is a 'white box' model, where the underlying biophysical descriptions considered to be the most important processes from the modeler's perspective are combined into an overall system model. The third category is a 'grey box,' with model elements that are determined from data only and elements that are determined from mechanistic descriptions.

1.3.1 Terminology

This work draws upon two somewhat separate fields, tree growth modeling and system identification, with somewhat varying terminology within and between them. It is worth noting upfront the definition of terms used here and other terms used in the literature (Table 1.2).

Table 1.2. Terminology used in this dissertation and synomous terms from the related literature.

Term used here	**Definition**	**Other terms used in the tree modeling and system identification literature**
Model	A mathematical description of relationships between observed quantities (Ljung 2011)	
Empirical model	Mathematical model of a system that does not directly reflect mechanistic processes, with parameters adjusted to fit the data	"black box model" Ljung 1999 "data driven model" Fan et al. 2015 "growth function" Sedmák and Scheer 2015 "inductive model" Young 2012 "statistical models" Weiskettel et al. 2011 "statistically based models" Pérez-Cruzado 2011
Process-based model	Model of a system and its behavior based on the submodels of the constituent processes that together determine the behavior and responses of the system (Landsberg 2003)	"first principles" Buck-Sorlin 2013 "hypothetico-deductive" Young 2012 "knowledge-driven model" Fan et al. 2015 "mechanistic" Monteith 1996 "theory-based" Jakeman et al. 2006 "white box" Ljung 2010
Hybrid model	A model which contains both empirical and process-based model features	"grey box" Ljung 1999 "knowledge and data driven model" Fan et al. 2015
System identification	Building mathematical models from observed input–output data	"inverse methods" Aster et al. 2011
Parameter	A model quantity that is determined by the modeler usually through observation and/or optimization	"constant" Monteith 1996 "input factor" Saltelli and Annoni 2010
True parameters	Parameter values that are accepted and used to generate reference outputs	"nominal parameters" Keesman 1989
Input	External signals that can be manipulated by the observer (Ljung 1999)	"driving variables" Landsberg 1981 "external force" Goudriaan and van Laar 1994 "external input" Schilders 2008 "primary factors" Pretzsch 2009

		"environmental variable" Fan et al. 2015
Output	Observable signals that are of interest (Ljung 1999)	"external output" Schilders 2008
Variable	A system quantity that is not an input, output, state, or parameter	"covariate" Weiskettel et al. 2011
Cost function	Function to be minimized with respect to model parameters	"misfit measure" Aster et al. 2011 "loss function" "objective function" Conn et al. 2000

1.3.2 Empirical models

For over 250 years, models built upon empirical growth observations have been used in forestry and are today still considered to be the best available models for yield estimation for specific species in specific locations (Pretzsch 2009, García 2011, Matthews et al. 2016). Empirical models are based upon statistical methods to fit growth equations to observed size data. These are static in the sense that they do not include time-varying inputs, and once parameterized give size predictions as a function of time only. Such models assume that a site has inherent productive capacity and that measures of past growth can be used to predict future growth for a given location, species, planting density, and silvicultural regime. Empirical models may be in deterministic form, which assumes the underlying growth processes are consistent over time, or in stochastic form, which emulates natural variation in the underlying growth potential (Vanclay 1994). Although such models are limited to the locations where the data was collected, empirical forest growth models have been widely used for operational planning and thus are considered a mainstay of forest modeling (Weiskittel et al. 2011). Their basis in data that are relatively simple to collect and accuracy in prediction have made empirical models the primary yield models

in forestry (Johnsen et al. 2001). Valentine and Mäkelä (2005) note that hundreds of empirical models for forest growth have been made.

Pretzsch (2009) reports that the earliest tree growth models were made for forest stand-level growth in the form

$$\text{Stand growth} = f(\theta, t) \tag{1.1}$$

where $f(\bullet)$is a sigmoid function of site fertility class θ and time t (Figure 1.1). The site fertility class parameter is a site- and species-specific measure of growth potential.

Early models were published in tabular form, showing predicted yields by tree age for a particular species and site class. Yield tables are based upon growth data collected for a certain site, planting density, and management regime. In many cases, these data have been collected for over a century, representing a large data set. It is commonly assumed that a monotypic (single species), even-aged stand of trees is present, rather than multiple species and uneven ages (Matthews et al. 2016). The first type of this model, which is still in use today, is based upon the concept of 'site index' first introduced by De Perthuis in 1788 (Batho and García 2006). Site index represents potential for growth on a certain site as determined by various measures including volume of standing timber, mean tree height, or top height (Pretzsch 2009). Timber volume is the estimate of interest, so estimated height is converted through an allometric (quantitative) relationship to a volume estimate. The primary model parameter is site index. These original models considered tree growth observations as the measure of site quality, rather than utilizing measurements of environmental drivers such as solar radiation, precipitation, and temperature.

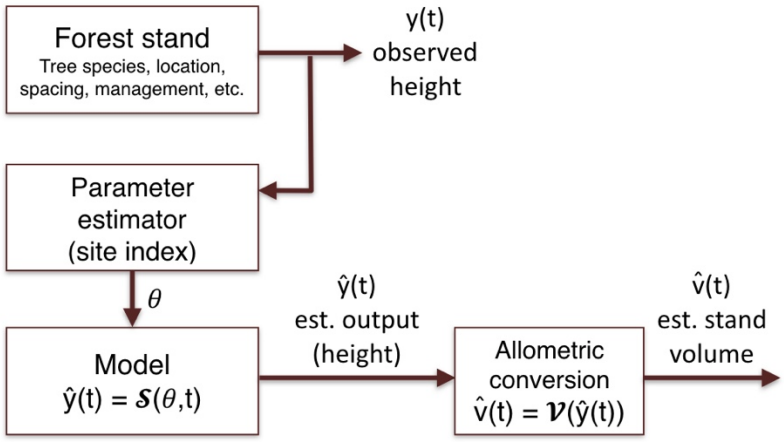

Figure 1.1: Early yield tables use observations of tree growth (usually height as shown here) to determine the site index θ (site fertility class), a parameter used in the model. Empirical models such as these do not explicitly include exogenous inputs.

Beyond the original empirical models built in terms of stand age and site index, more complex models that include silvicultural interventions such as thinning, pruning, fertilizing, as well as influences of improved germplasm and disease are now commonly used (Weiskittel et al. 2011, Weiskittel 2014). The model variables include tree species, initial DBH, tree height, height to crown base, spatial coordinates, site index, and topography (Table 1.3). Causal drivers or regulators of growth such as solar radiation, precipitation, and temperature, are implicit in the site-specific model parameters, as opposed to dynamical models where growth drivers are explicitly included in the mathematical descriptions.

Empirical models fit growth curves to data using regression analysis. The general form of the growth curves are sigmoidal, describing the opposing natural growth pattern of early exponential expansion followed by decline in growth rate. Zeide (1993) argues for the validity of sigmoid models for tree growth stating, "Biological growth, the outcome of numerous and enormously complex processes, appears remarkably simple, particularly for trees. As we combine more and more similar trees, the increase in their size follows an ever smoother sigmoid curve."

Further, Shvets and Zeide (1996) contend that the time-tested reliability of empirical models based on sigmoid growth equations accurately reflects the ecological processes of tree growth expansion and decline. In other words, even generic sigmoid growth equations can be seen as a high-level process-based model of macro-scale processes.

Table 1.3 Attributes of example modern empirical tree models by model type (after Weiskittel et al. 2011).

Model (source)	**Variables**	**Stand types**	**Variety of sampling designs**	**Silvicultural modifiers**
TASS	Tree species, DBH, HT, and HCB, spatial coordinates	Even-aged	No	No
SILVA	Tree species, DBH, HT, and HCB, spatial coordinates, and site characteristics (site index, ecoregion, elevation, slope and aspect)	All	No	No
PROGNAUS	Tree-list, geographical location, site fertility, and soil type	All	Yes	Yes
ORGANON	Tree-list and site index	All	Yes	Yes

Key: HT = tree height; HCB = height to crown base; DBH = diameter at breast height

An advantage of empirical models is the ability to select their form to reap multiple mathematical benefits in terms of their simplicity (or parsimoniousness) for mathematical analysis and calibration (identifiability of parameters) (Ljung 1999). Additionally, from a philosophical viewpoint, since data ultimately determine the validity of any model, empirical models have the advantage of being formed by inference from the available data. This inductive approach based on data and inference for empirical models may explain their continued usefulness

in predicting tree growth, despite decades of development of other types of models (Young 2012).

1.3.2.1 Limitations of empirical models

Empirical models are mathematical representations derived from growth measurements under specific site conditions for specific species and management (Johnsen et al. 2001). Their predictive ability has proven to be reliable for the same conditions for which they were developed. However, although empirical models are considered reliable in practice, they lack the ability to predict growth with changing site conditions (Valentine and Mäkelä 2005, Pretzsch 2009). Malézieux et al. (2009) state in support of large investments in models of biophysical descriptions of underlying processes, "Empirical models are useful for making predictions within the range of data used to parameterize them but are not suitable for extrapolation."

The success of empirical models at prediction implies that the average growth environment over many years is consistent, in other words, that natural year-to-year and decade-to-decade variations in weather over a period of years have time-invariant characteristics. This assumption has proven historically valid. However, as climate changes so do significant growth drivers such as solar radiation, temperature, rainfall, atmospheric CO_2, and occurrences of extreme weather events (USGCRP 2017). These changes cast doubt on the reliability of empirical models that were parameterized under now obsolescent operational conditions, as stated by Pretzsch (2009), "In the view of an increasing instability in forest growth due to atmospheric pollution, increased NO_x deposition, atmospheric CO_2 concentrations, and climate change, forest growth research cannot restrict itself

to statistical links between system variables at high levels of aggregation, but must understand and predict the responses of forest ecosystems."

Related to changing environmental conditions, empirical models work well in long-term predictions over averaged site conditions (Waterworth et al. 2007). However, since most empirical models operate on a 1- to 5-year time step (Pretzsch 2009), such models do not capture variability over short time periods. This may be especially significant in multispecies configurations where interspecific interactions drive short-term differences in growth dynamics. This limits the capacity of empirical models to predict short-term variations in growth that may be crucial for management decisions in intensively managed systems (Fontes et al. 2010, Pérez-Cruzado et al. 2011).

1.3.3 Process-based models

The limitations of applying empirical models to novel situations gave rise to ever more complex models based on descriptions of the underlying processes of growth. Models known as 'process-based' describe growth phenomena as a function of first principles or mechanistic knowledge of internal dynamics (Landsberg 1981, 2003a; Whisler et al. 1986; Goudriaan and Van Laar 1994; Weiskittel et al. 2011; Buck-Sorlin 2013). While empirical modeling describes systems behavior based upon observed relationships, process-based models describe the underlying structural and functional relationships of growth. Korzukhin et al. (1996) describe a natural progression from empirical to process-based models, "Empirical models have their greatest value during the descriptive stage of knowledge development, and process models have their greatest value during the explanatory stage." This progression to process-based models required computational capacity to simulate complex mathematical representations. The

rise of process-based models in the early 1960's coincided with the availability of mainframe computers, and accelerated with the convenience of desktop computers in the 1980's (Monteith 1996, Priesack and Gayler 2009, Pretzsch 2009, Yin and Struik 2010).

Process-based models include differential or difference equations describing complex interconnected processes in response to inputs (growth drivers). In this mechanistic framework, plant growth is seen as the conversion of resources such as sunlight and water into biomass (Monteith 1972, Black and Ong 2000). The processes are biophysical (physics applied to biological processes) and the inputs are environmental such as solar radiation, precipitation, and temperature. In application, process-based models consist of the mathematical equations, parameters embedded in those equations, links between equations, and a computer implementation of the model (Vanclay 1994).

Rather than identifying a model structure and parameters from observations of the output measurement of interest (e.g., tree biomass), process-based tree growth models utilize descriptions of underlying biophysical processes to construct dynamic descriptions of how input drivers influence growth. This approach is rooted in the modeler's confidence in their knowledge of the underlying mechanisms in the system (Young 1983, Landsberg 2003b). It also reflects a shift of focus from prediction to investigation of underlying growth mechanisms (Pretzsch 2009).

Typically, process-based models are constructed from submodels, each of which describes part of the underlying system dynamics that are of interest, and each having input and output signals connected to other blocks in the model. Each submodel is investigated and parameterized via detailed experimentation under

controlled conditions in the field or laboratory. Submodels are often nonlinear in the input-output relationships and in the parameters. The modeling process consists of connecting descriptive submodels based on prior knowledge and deductive reasoning in a way that the resulting full system model is believed to mimic the system's physical nature. This "hypothetico-deductive" modeling philosophy (Popper 2005) is widely used in modeling natural processes (Young 2013).

Ljung (1999) explains this model building strategy, "One route is to split up the system, figuratively speaking, into subsystems, whose properties are well understood from previous experience. This basically means that we rely on earlier empirical work. These subsystems are then joined mathematically and a model of the whole system is obtained. This route is known as *modeling* and does not necessarily involve any experimentation on the actual system." In the resulting full model, interactions of the submodels may give rise to unexpected behaviors that suggest further model revisions or new realizations about system behavior. This process of building complex system models from submodels and refining them is a valuable heuristic process, whereby the underlying processes and their role in the whole system are better understood (Yin and Struik 2010).

For tree growth, various research perspectives give rise to numerous potential submodels. As an example, Jarvis (1981) proposes 14 interconnected submodels (see Table 1.4), which can be described mathematically. Submodels can be interconnected in various ways (see Figure 1.2). Process-based models for tree growth have been made on many different scales in plant biology (molecular, cellular, organism), plant interactions (monoculture, polyculture, uneven age populations), and physical size (tree, forest, and landscape levels) (Pretzsch 2009

pp. 27 and 30; Landsberg 2003a,b). The interests of the modeler regarding scale have understandably led to numerous types of models of varying complexity. In addition to physical scale, the model's temporal resolution is determined by the system dynamics of interest. Depending on the time domain of the model dynamics and inputs/outputs of interest, data sampling may be required with minute, hour, day, or longer time step (Goudriaan and van Laar 1994).

Table 1.4. Jarvis (1981) presents this table showing how various submodels applicable to tree growth are connected to each other. For example, in this system representation, the soil and plant water submodel generates outputs that influence phenology, photosynthesis, soil and plant nutrients, and fine root growth. There are 29 possible output-input combinations between submodels, plus the "weather variables" that are inputs to 13 submodels. One can see from this example that the interactions between submodels can become quite complex in process-based models.

Submodel giving output	**Submodel receiving input**													
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
1. Phenology	O	×				×								×
2. Light/Radiation		O	×	×	×			×						
3. Conductance			O	×	×									
4. Photosynthesis				O		×								
5. Transpiration					O				×					
6. Allocation	×					O	×				×	×	×	
7. Respiration							O							
8. Evaporation								O	×					
9. Soil & plant water	×			×					O	×		×		
10. Soil & plant nutrients	×		×							O				
11. Stem/branch growth		×			×						O			×
12. Fine root growth					×							O		
13. Coarse root growth					×								O	
14. Population			×											O
15. Weather variables	×	×	×	×	×	×	×	×	×	×	×	×	×	

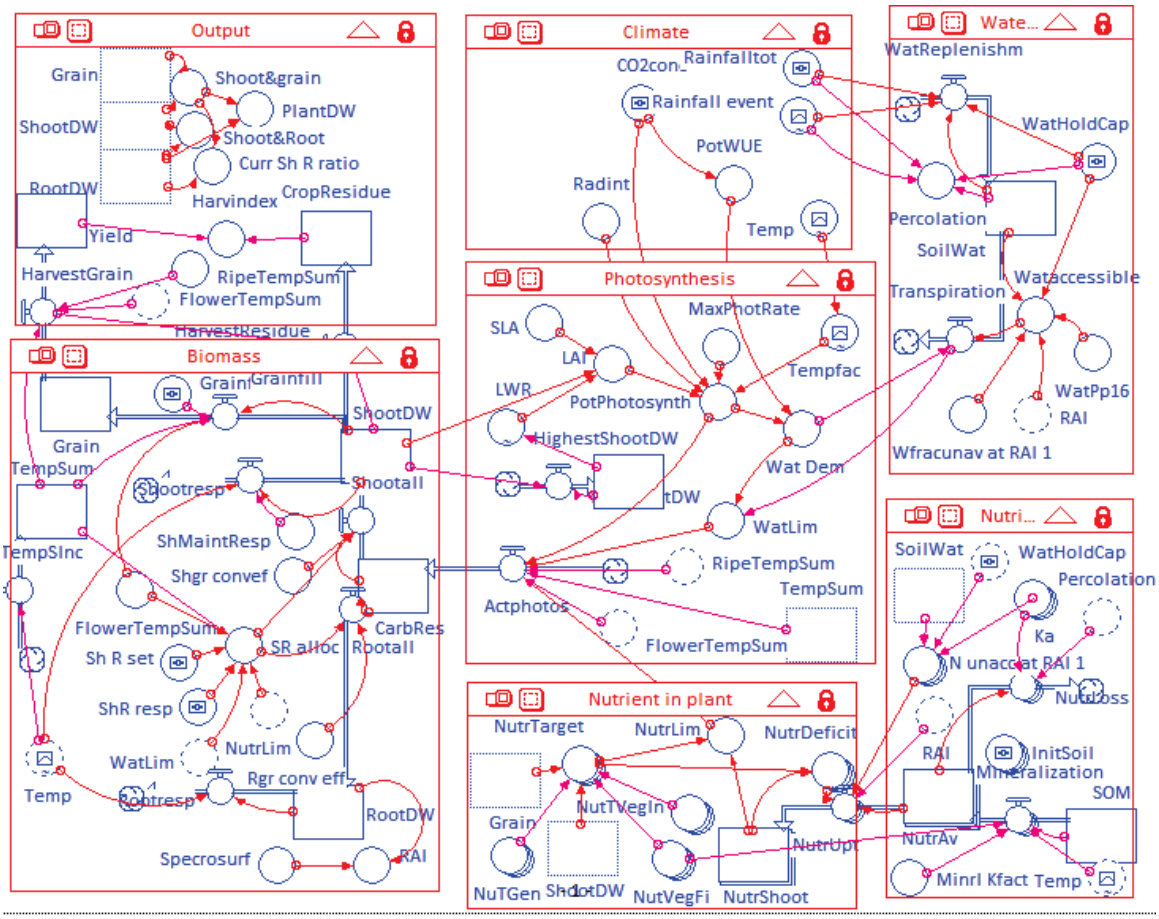

Figure 1.2. van Noordwijk et al. (2011) presents this example process-based plant growth model. The inputs are sunlight, rainfall, and temperature, and the outputs are daily biomass development of leaves, roots, flowers, and seeds. Submodels include conversion of sunlight to biomass, biomass allocation in the plant, nutrient balance in the plant and soil, and soil water. Note the numerous connections—input-output signals—between submodel boxes.

1.3.3.1 Limitations of process-based models

There is a general consensus in the tree growth modeling community that empirical models are well suited for growth predictions, while process-based models are best suited for understanding the underlying mechanisms of growth. From a survey of process-based forestry and agroforestry models, Malézieux et al. (2009) conclude, "Most of the [existing process-based] models are used as research tools rather than management tools." Observing that relatively simple empirical model based on sigmoid growth curves are used now more than ever in predictive

tree growth modeling, Sedmák and Scheer (2015) observe, "A direct prediction of the integral quantities of growth at individual or population level by means of simpler growth functions can be more precise than prediction from the sum of the individual components of growth (cells, tissues and organs) obtained by more complicated process models." Regarding the value of process-based models for prediction, Mäkelä et al. (2000) assert, "Process models are considered to embody too many uncertainties and to require too many poorly known parameters for their projections to be as reliable in practice as those of empirical models."

Even though process-based models are constructed from independently parameterized submodels, the input-output dynamics of the whole model must be validated and parameter values calibrated. Noting the problem of complexity in process-based models, Young (2011) says, "In such examples, particularly those concerned with 'natural' systems, such as the environment, climate and the economy, the model contains a surplus content that often owes more to the perception of the scientist than to its evaluation against observational data."

Submodels are often assumed to be well tested, credible, and validated models for underlying dynamic processes. However, by connecting submodels into a large model, complex feedback loops are often created within the model, possibly altering the dynamic behavior of submodels in ways that cannot be predicted. As Yin and Struik (2010) state, "The rules by which the elements or processes interact give rise to systems behaviour and emerging properties, which may well be unexpected and even counterintuitive." Due to model complexity, it may be very difficult or impossible to determine the underlying cause of modeling errors in the full system output (Priesack and Gayler 2009). Schmid et al. (2006) conclude in a survey of forest management models that "more complex models ...

are sensitive to uncertainties in structure, parameter values and input data." Further, certain submodel parameters may be problematic to identify from experimental data, potentially leading to large parameter errors that in turn may generate large output errors (Beck 1983, p. 6). Mason et al. (2007) note that parameter errors propagate due to the recursive nature of models. Therefore, a strategy that limits the model complexity and number of parameters—the least complex model that adequately reflects system behavior—is preferred (Ljung 1999, Young 2011) (see Section 1.4.3).

Although modelers use model error as an opportunity to further refine the model by modifying representation of internal processes and how submodels are interconnected in the model, others have found this to be a fundamental flaw in the modeling process. Sinclaire and Seligman (1996) argue in a review of the value of process-based crop models, "The fundamental difficulty is that all [process-based] models are basically a collection of hypotheses and not a single falsifiable hypothesis, so they inherently cannot be validated (Pease and Bull 1992; Oreskes et al. 1994). Not only can other collections of hypotheses approximate the experimental results equally well, but the validation data themselves are flawed by substantial experimental and observational error." They conclude that process-based models have more value as heuristic than predictive tools.

There may be ample data from isolated experiments to parameterize submodels with confidence, but inadequate data for validation of the whole model. Often, there are fewer data points than the number of model parameters (Keesman et al. 2011). In such cases, validation may consist of a qualitative comparison of model output with available data and judging whether the predictions are "reasonable"

(e.g., Graves et al. 2010). This qualitative evaluation may take place after making adjustments to the model structure and parameters based on quantitative analysis and guidance in the form of expert knowledge (Keesman et al. 2011). As Passioura contends in his 1973 critique of the process-based model building strategy, "To test the gross output of the model is easily done; to thoroughly test the internal behaviour is, in practice, virtually impossible. But the test on gross output is very feeble. It is almost always unsuccessful, and the modeller is tempted to fiddle the parameters in his model until it does work with offending data. If this fiddling does not work, he creates a new subsystem or new relations between some of the old subsystems."

1.3.4 Hybrid models

In an early system identification paper, Bellman and Åström (1970) state, "[the] goal in many identification problems, in both industry and the biosciences, is to combine *a priori* knowledge with experimental data." Hybrid models combine aspects of empirical and process-based models to take advantage of each approach's strengths: the relatively reliable predictive abilities of empirical models and the ability of process-based models to incorporate environmental drivers and physically interpretable parameters for increased site specificity (Fontes et al. 2010, Weiskettel et al. 2011). Young (1983) casts doubt on the use of purely mechanistic models for complex natural systems, suggesting the addition of an empirical model component, "the size and complexity of many natural systems, such as those encountered in environmental and economic research, are such that the mechanisms governing the change in the observed system variables and their interrelationships are rarely fully understood *a priori*."

There exists no general method for combining the two modeling approaches, rather, combinations of the approaches are problem-dependent and rely upon the modeler's confidence in the problem domain knowledge (Fan et al. 2015). As a result, hybrid modeling opens a wide array of potential modeling strategies. Landsberg (2003b) notes regarding the potential for hybrid models, "Combining the advantages of process-based and conventional empirical models will be the management tools of the foreseeable future." In a survey of tree and forest models Pretzsch et al. (2015) notes that of the 54 models considered, 10 are classified as hybrid models, with process-based models dominating the field. Examining the virtues of the hybrid approach, Valentine and Mäkelä (2005) argue that while the advantages of combining empirical and process-based models are open for discussion, "any movement toward a common model seems sure to increase communication between the two schools of thought." Further, Valentine and Mäkelä (2005) suggest that hybrid models are an opportunity for empirical modelers to transition toward process-based descriptions.

Empirical and process-based approaches are usually treated as distinct in the literature, however, they can also be seen as a continuum (Korzukhin et al. 1996; Weiskittel et al. 2011, p. 2). Although empirical models rely primarily on observational data for curve fitting, process-based models also use observational data in formulating submodels of the underlying mechanistic processes, to determine process description parameters, and for model validation. Conversely, empirical models also describe a type of process, such as the increase in timber volume in a forest stand as a function of age, even though they do not explicitly describe the underlying growth mechanisms (Korzukhin et al. 1996; Shvets and Zeide 1996). By virtue of the designation "hybrid," such models explicitly combine

elements from both empirical and process-based strategies (Waterworth et al. 2007).

Because of the extremely wide range of possible approaches to hybrid modeling, generalizations are difficult to make. For forest growth modeling, Weiskettel et al. (2011) summarizes the variety of existing hybrid models in three classes

1. Statistical growth equations with a physiologically derived covariate [explanatory variable].
2. Statistical equations with a physiologically derived external modifier.
3. Allometric models.

1.3.4.1 Limitations of hybrid models

Hybrid models require *a priori* mechanistic knowledge of the system, just as with process-based models. High quality data is also needed for the empirical portion. As noted by Pretzsch (2009) regarding forest growth hybrid models, "For a sound performance, the initial values and calibration of the internal estimator functions [empirical model portions] of such models must be backed by a regionally extensive network of inventory data on biomass, soil conditions, climate, and growth." Additionally, whole system data is also required for calibration and validation, as it is for any model. One could say in general that the knowledge and data requirements are greater for hybrid models than either process-based or empirical models.

The process-based portion of the model potentially has the same pitfalls as process-based-only models regarding complexity and parameter identifiability. Further, there is no standard way to integrate the empirically based portion of the model with the process-based portion (e.g., Fan et al. 2015), meaning that

there is no generic formal way to determine the validity of the whole model, just as with process-based models, without adequate observational data.

Weiskettel et al. (2011) suggest that where submodel dynamics are replaced by empirically derived parameters, model output can be highly sensitive to errors in such parameters, noting that, "The limitations of these ... [hybrid] approaches are that they are insensitive to important site factors and cannot represent the influence of variable weather patterns." Also, Weiskettel et al. (2011) raises questions of model complexity and parameter identifiability in that the relatively simple hybrid model 3-PG, "is not path invariant and is highly recursive, which means that the model can be fitted to data in a variety of ways and errors may propagate when dependent variables from one month are used as independent variables during the next month." In other words, hybrid models share the limitations of both empirical and process-based models. A comparison of the strengths and limitations of the three model classes are given in Table 1.5.

1.4 A system identification approach

System identification builds mathematical models based upon observed system input-output data. A well-known and straightforward definition for system identification by Zadeh (1962) is often quoted, "the determination, on the basis of input and output, of a system within a specified class of systems, to which the system under test is equivalent." In other words, system identification is about building models based on input-output data.

Table 1.5: Comparison of tree growth modeling approaches (after Weiskittel et al. 2011)

Model type	**Important uses**	**Strengths**	**Limitations**
Empirical	Predict yields; Update forest inventories; compare forest silvicultural treatments; estimate sustainable harvests	Robust; long history of development; rely on data generally available; output geared for operational decisions; can select well behaved *a priori* mathematical expressions	Requires high quality data; can extrapolate poorly to new environments; generally insensitive to climate
Process-based	Understand the underlying Mechanisms influencing growth; test hypotheses about plant behavior; predict potential forest productivity	Can theoretically extrapolate to novel situations; sensitive to climate; mechanistic and parameter interpretation; assist with heuristic learning	Dependent upon difficult-to-measure parameters; data not widely available; high computational demand; tendency to high complexity; output often unusable for operational decisions
Hybrid	Predict growth using climatic factors; prediction of novel forest silvicultural treatments; address shortcomings of empirical and process-based models	Sensitive to climate; reduce the number of required parameters; may be able to use traditional forest inventory data	Accuracy improvements can be minimal when compared to a purely statistical approach; climate and soils input data not widely available

System identification emerged as a foundation of feedback control engineering, and thus has a rich representation in the control literature since the mid-1960's (Balakrishnan and Peterka 1969, Åström and Eykhoff 1971, Eykhoff 1981, Ljung 1996). Considering the current role of system identification to be underutilized by modeling communities in other fields, Ljung (2010) states, "Even though applications of system identification apparently go far beyond the automatic control community, ... the impact of system identification theory on other research areas seems modest."

System identification begins with data, which, together with *a priori* system knowledge are used to choose candidate models that are parameterized based on data, followed by model evaluation. If the model is deemed inadequate, the process is repeated with new experiments and/or revised models, and so forth, resulting in an identification loop (see Figure 1.3) (Ljung 1999; Keesman 2011).

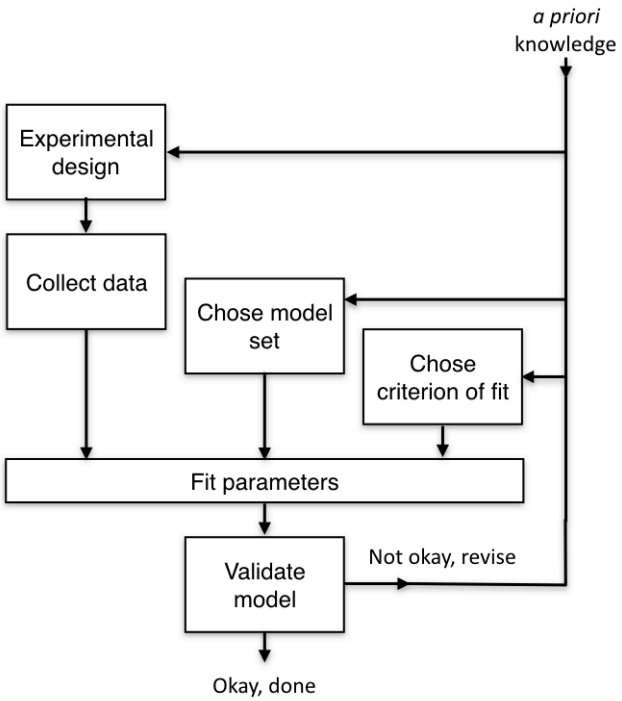

Figure 1.3 A representation of the system identification process. This process assumes that observation data are available for model calibration and validation. (Figure from Ljung 1999).

1.4.1 Resistance to system identification in tree growth modeling

A system identification approach is largely overlooked in tree growth modeling due to various considerations, including

- Lack of input-output observational data for the system under consideration.
- Attachment to a mechanistic model interpretation.

- Attachment to physical interpretation of parameters.

Since the 1970's there has been a strong underlying assumption that mechanistic, process-based models (including hybrid) are the way forward for predictive tree growth modeling. In 1981 Landsberg stated regarding models that can be used for management decisions, "Models intended for use as management tools are likely to be more useful and reliable if they are based on, and incorporate, the mechanisms underlying observed and simulated responses." There continues to be a broad consensus that process-based and hybrid models will continue to dominate predictive tree growth modeling for both research into underlying processes and prediction for planning and management purposes (Malezieux et al. 2009, Pretzsch 2009, Weiskittel et al. 2011).

However, as outlined in Section 1.3, process-based and hybrid models have a number of interconnected limitations including high complexity leading to difficulties with parameter identification and model validation. Where studying the underlying biological processes is not of primary interest, but prediction of growth measures is the primary purpose, process-based models for tree growth are promising candidates for model complexity reduction [reduced-order parameter estimation] from measured input-output data in a system identification context. The system identification approach is particularly relevant in practice, where model states or internal model signals are not easily measurable, or their observation is impractical or expensive (Ljung and Glad 1994b). Investigation of model reduction in plant growth is particularly timely, as numerous institutions and governments throughout the world are seeking innovative approaches to predictively model forest/agroforestry productivity under changing climate conditions (e.g., AFRI 2016).

In an early paper on system identification of process dynamics, Godfrey and Brown (1979) state that, "*A priori* knowledge, as embodied in a model structure, is essential for effective identification." Since process-based biological growth models include process knowledge in their mathematical structure, this suggests that such models are an interesting starting point for a system identification approach (Ljung 2010).

Despite the lack of formal system identification approaches in tree growth modeling, a case can be made that a system identification approach relates well to empirical, process-based, and intermediate (hybrid) tree growth modeling. A comparison of empirical and process-based models with a system identification approach is given in Table 1.6. Further, precedent in addressing the three considerations noted above using system identification approaches can be found in the tree growth modeling literature. It is instructive to examine each of the three considerations about system identification in further detail.

Table 1.6. Comparison of model approaches

	Parameter determination	**Equations**	**Validation**
Empirical	Curve fitting from data, physical interpretation	General, black box growth equation	Input-output data
Process-based	Submodel studies, physical interpretation, adjustments in full model	Mechanistic descriptions	Input-output data for submodels and whole system
System identification	From measurable system input-output data, may not have physical interpretation	General growth equation, mechanistic description, or hybrid	Input-output data for whole system

1.4.1.1 Consideration #1: Lack of data

In a statement that applies to all models, Weiskittel et al. (2011) assert, "Models are often only as good as the data used to construct them." As noted in Section 1.3.3, process-based models are often the only means forward in modeling due to the lack of data. In this sense, process-based models serve as a stand-in strategy for little-studied natural systems, an approach that is widely accepted in the literature (Young 2012). When data are sparse or non-existent, the modeler relies on physical insights into the underlying processes and intuition to devise a mathematical description, often with little regard to model complexity. However, Beck (1983) suggests that the lack of available data is a factor that encourages increasingly complex models, with the prevailing assumption being "more detail necessarily means a better model."

While there exist large empirical databases of tree growth data for outputs such as height (Pretzsch 2009, Weiskittel et al. 2011), data for the inputs of interest are often missing. Also, input-output data must be collected at the temporal resolution suitable for the process dynamics that are of interest. Much of the legacy data for tree growth was collected at 5-year intervals (Pretzsch 2009), in effect missing important process dynamics that may occur within weeks, days, or months (Goudriaan and van Laar 1994). Therefore, rather than a blanket statement about the lack of data, it is more accurate to say that there is a lack of adequate data appropriate for system identification purposes.

From a process-based modeler's perspective, one might associate 'sufficient data' with collection of input-output observations for the system outputs as well as all associated submodels, which, of course, would be a highly impractical and expensive undertaking. However, a system identification approach requires only

observation of inputs and outputs for the whole system. In the case of tree models, relevant inputs include environmental drivers of growth such as solar radiation, temperature, precipitation, atmospheric CO_2 concentration, and soil nutrient status, while relevant outputs usually include measures of size. Most of these data can now readily be collected and sent to computer databases automatically and cheaply (see Chapter 5), suggesting that data collection within a system identification framework can be relatively straightforward (although it still requires years of observation).

As the choice of signals to be measured and the sampling interval are integral to model development, it makes sense to thoroughly evaluate candidate process-based models prior to initiating field studies. For such analysis, the model candidate is assumed to be a surrogate for the true system, in other words, that the input-output data generated by model simulations accurately represent the true system (Young 2012). Using this strategy, the dearth of experimental data is overcome by an abundance of simulation data to investigate important model properties before experimental fieldwork takes place.

1.4.1.2 Consideration #2: Bias toward mechanistic interpretation

When referring to development of predictive models for grower usage, Whistler et al. (1986) state, "If these models are comprehensive and mechanistic, as we believe they should be, then they will allow the grower to use management 'what if' games to check on the profitability of many options." This statement exemplifies the commonly held belief that if process-based models faithfully capture causal dynamics, then their application can be extrapolated to a wide range of environments. Therefore, an approach that prioritizes accurate input-output behavior over the accurate representation of internal mechanisms that are

perceived to be important by the modeler, might not be accepted for philosophical reasons (even if the resulting model is shown to be valid and credible).

As Young (1983) remarks on acceptance of models that are outside the norm, "If the scientific establishment is firmly committed to a particular type of model for a physical system, then it may at first be difficult for the systems analyst to gain acceptance and credibility for a less conventional representation, even if the model-building procedure has been rigorous and is seen to conform with the basic tenets of the scientific method." And further, on developing reduced complexity models, Young (2013) remarks, "Consequently, the resulting model may not always be fully acceptable or credible to an audience that has been educated to believe strongly in hypothetico-deductive [process-based] modeling based on conceptual, often deterministic, simulation models."

However, in characterizing the feasible parameter set for non-linear growth models, Keesman and Stappers (2004) note, "When dealing with complex simulation models for which the internal structure is too complicated to be analyzed analytically via e.g. linearization or interval analysis, we have to rely on the input-output behavior of the model." Our approach here retains mechanistic interpretation of process-based models, but is based upon input-output behavior of the model. The reduced-order parameter optimization may lead to parameter estimates that no longer have physical interpretations. Such a strategy is in contrast to black box models that are used as a vehicle to fit the data, but that lack any apparent mechanistic interpretation.

1.4.1.3 Consideration #3: Bias toward parameters having physical interpretation

In process-based modeling, parameters are intimately tied to the mechanistic process description, meaning they also carry a biophysical interpretation such as physiological rates or allometric relationships. Parameter values are usually determined for submodels through controlled experimentation. Where parameters may not be easy or possible to measure even through controlled studies, they are obtained from previous related studies in the literature or extrapolated by experts from previous reputable scientific study. Examples of important parameters with physical interpretation used in tree models include the efficiency of conversion of intercepted solar radiation into biomass (Monteith 1972), the extinction coefficient of light through the canopy (from the Beer-Lambert Law for light attenuation), and water transpiration per unit of biomass (Landsberg and Sands 2011a). The physical interpretation of parameters can aid in model analysis. Additionally, parameter ranges allow for comparisons of different species and configurations. A physical interpretation can also aid in determining parameter values through expert recommendations or in constraining values in parameter optimization (Keesman et al. 2011).

This ideal scenario of determining physically interpretable parameters in process-based models is often not possible in practice. As Landsberg and Sands (2011b) state on the feasibility of establishing physically interpretable model parameters, "Lack of data, poor understanding of processes, or problems with scaling-up, may make it necessary to estimate the parameters of sub-models using empirical data relating to the whole system. This is frequently achieved by adjusting parameter values to improve the fit between model output and observations at the level of the whole system." The idea of using observations at the level of the whole

system rather than for sub-processes is closely related to the system identification approach that will be used in this analysis. This concept is central to the hybrid models described in Section 1.3.4, where whole system observational data are used to fit parameters in parts of the model. Such fitted parameters may lose their physical interpretation, and are seen only as a means to adjust model fit (Ljung 1999).

As an example, Valentine and Mäkelä (2005) acknowledge a shift away from physically interpretable parameters in developing a hybrid model for tree growth, "In principle, the values of ... parameters may be estimated by lower-level process models. Alternatively, the physiological and morphological parameters combine, under reasonable assumptions, into a set of aggregate parameters, whose values can be estimated from inventory data with a statistical fitting procedure." In other words, the individual parameters may lose their physical interpretation through the fitting to empirical data, but the suite of identified parameters combine into aggregate parameters that are used to more accurately model the overall system dynamics. Such shifts in the interpretation of parameter meaning show that an opening already exists in the tree growth modeling community to this system identification approach to parameter interpretation.

In an analysis of 16 biological models, Gutenkunst et al. (2007) show "that predictions from most models will be very fragile to single uncertain parameters and that collective parameters fits can often yield tight predictions with loose parameters." In other words, they argue for the superior value of fitting several important parameters ("collective fit") based on model predictions, rather than on the value of individual parameters.

1.4.2 Why consider system identification for tree growth modeling?

The shortcomings of process-based models highlighted in Sections 1.3.3.1 suggest excessive complexity and accompanying over-parametrization are central issues to be investigated during model development. This leads to the following fundamental questions about the use of process-based models for prediction, which form the basis of this research

1. Can the parameters be identified in the full model from input-output data?
2. How many model parameters need to be estimated in order to achieve an acceptable fit to data?
3. How can one determine which parameters are important to estimate and which can be set to nominal values?

In the case of process-based models for systems where data are scarce or non-existent (such as in modeling agroforestry systems), the models are built from submodels based upon prior system knowledge and modeler insights into the dynamics of interest (Ljung 1999, Young 2012). This leaves a critical gap in the classic system identification process where the data are absent. For tree growth, how can the model be tested against data without running expensive, long duration field experiments?

The strategy here is to evaluate a candidate process-based model for parameter identifiability and complexity before running expensive field tests. One could see this process as a stand-in for the evaluation of model behavior that has been done for a range of black box models used in system identification (e.g., linear regression, neural networks) with the emphasis on questions 1–3 above. This

approach can be seen not only as model building, but the preparation for experimental design, as depicted in Figure 1.4.

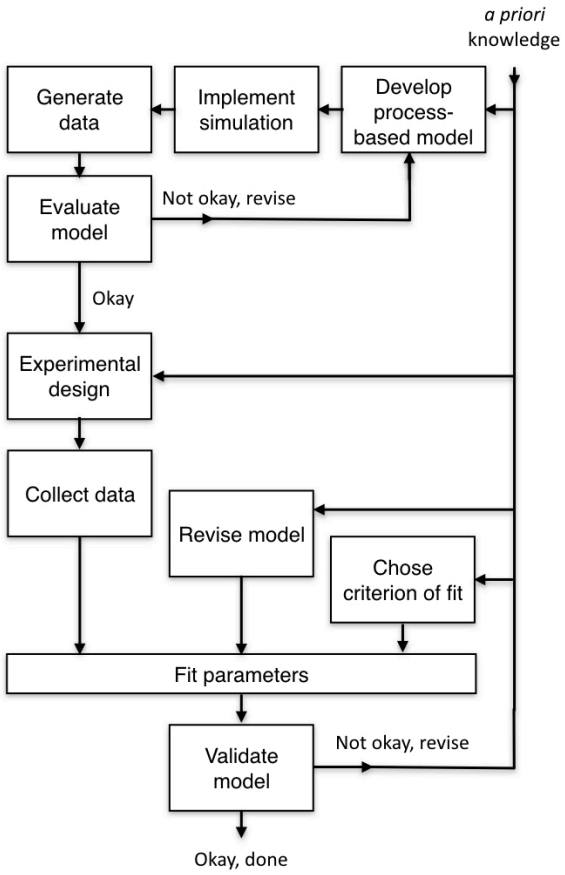

Figure 1.4 A revision of the system identification loop presented by Ljung (1999), where the process-based model is constructed and evaluated prior to field experimentation.

To summarize, the following assumptions lead to a system identification framework

- Output prediction is of primary importance (rather than the understanding of internal mechanisms).
- Signals within the model are no longer necessarily assumed to have physical interpretations.
- Parameter interpretations may no longer be valid.

Given these relaxations in a process-based perspective, system identification approaches can be used to give insights into model complexity, parameter identifiability, and model validity.

1.4.3 Model complexity

William of Occam's principle of parsimony (aka Occam's Razor) states, "entities should not be multiplied unnecessarily," which is widely interpreted as, "given two explanations of the data, all other things being equal, the simpler explanation is preferable" (Blumer et al. 1986). While there is controversy over whether the simplest explanation of past observations is the always the best predictor of future observations (e.g., Domingos 1999), it is generally accepted that a simpler explanatory hypothesis is preferred. As Boote et al. (1996) concisely state, "Simple crop growth models are easy to comprehend, often require fewer inputs, and often are easier to use and apply."

While many areas of systems modeling have progressed from mechanistic descriptions toward input-output (I/O) identification of efficiently parameterized models with parsimonious form, modern tree growth models are still built primarily upon process-based descriptions (Fontes et al. 2010, Weiskittel et al. 2011). Such models contain complex nonlinear structures and often dozens or hundreds of parameters to describe underlying growth processes such as photosynthesis, biomass partitioning, light interception, respiration, and carbon allocation (Johnsen et al. 2001, Landsberg and Sands 2011b, van Noordwijk et al. 2011, Heuvelink and de Reffye 2015). Despite the resulting model complexity, this approach is considered justified by the plausibility of the underlying process submodels and their connections within the model. However, model complexity reduction is often not attempted for process-based plant growth models and in

other biological and natural resources fields (Jakeman et al. 2006, Crout et al. 2009).

An underlying concern about complex process-based models lingers, as expressed by Beck (1983, p. 7), "... it may be that the computer era has merely fostered the growth and popularity of large simulation models with little accompanying increase in understanding." As stated by Weiskittel et al. (2011, p. 228), "Some important drawbacks to process models are that they are often quite complicated with output that is of little practical interest, which makes them difficult to parameterize or even operate due to high input data requirements." Awareness that the complexity of natural growth processes can lead to models of varying levels of complexity raises the notion that the model should fulfill the objectives of the modeler (Jakeman et al. 2006). A process-based model that is designed for study of underlying growth processes may be very different from a model that is meant to accurately predict outcomes. Regarding the ever growing complexity of process-based ecological models, Beven (2002) says as quoted by Sivakumar (2008), "The development of more and more complex models that incorporate more and more detail about processes, but which introduce more and more parameters that must be calibrated, does not appear to be the future ... The future of (environmental) modeling will have to place more emphasis on the value of data, carefully collected for specific purposes, and on parametrically simple robust models, carefully designed for specific purposes."

In such complex models, numerous parameters require *a priori* knowledge or are derived from often quite expensive and time-consuming field study (e.g., maximum canopy quantum efficiency, respiration rate). While many parameters are assumed to be the same in practice for a wide variety of trees, there is

controversy about how they might change for different species or across widely differing conditions (Landsberg and Sands 2011, p. 236; Reichert and Omlin 1997, p. 290). Additionally, it is often unclear how small errors in the parameters impact fitting of other parameters, as observed by MacLean et al. (2012) "Often, model parameters are highly correlated, and a change in one parameter can be almost completely compensated by changes in the other parameter values. In these cases, the parameter estimates cannot be considered reliable." Mäkelä et al. (2000) state in their analysis of practical implementation, "Process models are considered to embody too many uncertainties and to require too many poorly known parameters for their projections to be as reliable in practice as those of empirical models."

Researchers are addressing the complexity of process-based models with various approaches to simplification. One approach is to build process-based models from the ground up with an eye toward parsimonious descriptions and relationships. Recognizing the particular pitfalls of process-based models where quantitative long-term data are lacking in agroforestry, the Yield-SAFE model was built using a 'minimal modeling approach' (van der Werf 2006). Despite this conscious modeling strategy, determining parameters in the resulting model was challenging, requiring both quantitative analysis and expert judgement (Graves et al. 2010, Keesman et al. 2011). The Yield-SAFE model will be used as a test case for a proposed reduced-order parameter estimation procedure presented below.

1.5 Example model reduction techniques

In order to reduce model complexity and address a range of associated complexity issues noted in Section 1.4.3, many model reduction schemes have

been developed. In general, model reduction methods approximate the most important dynamics of the system by considering the relationship between inputs and outputs (Schilders 2008). This means that model reduction is closely related to system identification. Methods for reducing the complexity of linear models are well-studied and straightforward, less straightforward for linear approximations of nonlinear models, and most challenging for reducing nonlinear models to simpler nonlinear models (Ljung 2010).

General model reduction techniques such as Proper Orthogonal Decomposition (POD) project high dimensional data onto a lower dimensional space to derive a low dimension models that captures most of the dynamics of the full model (Hinze and Volkwein 2005, Kerschen et al. 2005, Schilders 2008). Although POD can be applied to nonlinear models, it yields an approximating linear manifold, which may be considered a serious limitation in its application to nonlinear models (Kerschen et al. 2005).

Recognizing that many process-based models are "overparameterised [and] may have poor predictive performance," Cox et al. (2006) followed by Crout et al. (2009) and others, approach model reduction by setting signals in the model to fixed nominal values and evaluating how well each model version performs. The signals considered are chosen subjectively, but generally have a mechanistic interpretation. An exhaustive search of all possible combinations of signal replacements with constant values are assessed in order to determine which model dynamics might be excluded from the model. Essentially, this method answers the question, "Which process submodel dynamics can reasonably be neglected?" during the model development process (Cox et al. 2006).

Also based upon analysis of dynamic modes in process-based models, Young and Ratto (2009) present a seven-phase strategy for reducing the complexity of process-based models termed Data-Based Mechanistic (DBM) modeling. After implementing a computational version of the model, the process begins with stochastic analysis using Monte-Carlo simulations to determine the relative importance of different submodels in explaining the dominant model behavior. From this step, a process called "dominant mode analysis" is performed over a user-defined parameter range to determine a low order approximation of the original model. This process generates a mapping of the unreduced model to a reduced order model called the Dynamic Emulation Model (DEM) that can replace the full model over a range of parameter values (Young 2011 p. 360). All of the analysis is done in the absence of experimental data, with the original model serving as a surrogate for the 'true system.'

Regularization is a commonly used technique to determine which parameters are more important than others by adding a term to the cost (aka loss or performance) function that penalizes the magnitude of the parameter vector (Sjöberg et al. 1995). The cost function is typically the sum of the squared errors of the output sequences, while the regularization term is proportional to the summed squared parameter values. By setting the relative weighting of the two terms, parameters that strongly influence the cost function will tend to be estimated, while less important parameters will be 'pushed' toward zero by the penalty term. The parameter estimates can also be kept close to their *a priori* estimates by altering the penalty term and using a weighting matrix (Ljung 2018).

1.6 Model validation

Model validation has several meanings, including testing against real data (Monteith 1996) or a phase of rigorous analysis where attempts to falsify a model as a faithful representation of the real-life system fail (Young 1983). The research presented here should be seen as a small part of the overarching process of confirming model validity (Jakeman et al. 2006). As illustrated in Figure 1.4, the work of parameter estimation can take place prior to field studies, and, in fact, can inform experimental design by determining which data are needed in order to assure a good model fit.

Therefore, for our purposes, Ljung's (2010) definition of validation is used, "The process of ensuring that the model is useful not only for the estimation data, but also for other data sets of interest. Data sets for this purpose are called validation data." One may be able to adjust parameters well for one set of input-output data, but the same parameter settings may give a poor match to other input-input data from the same system. In this *a priori* analysis, the model with accepted parameters serves as a surrogate for the true system, allowing us to use its input-output sequences as reference data. For model validation, the model is driven by various input sequences from the same location as those used for parameter optimization and compare model fit with optimized parameters (see Section 3.1.3).

1.7 Hessian analysis

As will be seen in the examples of Chapter 3, the cost function Hessian is a function of the location in the parameter space (parameter settings), the input(s),

and the model output(s). Since the model output(s) are also functions of the parameter settings and input(s), the Hessian can be expressed as

$$H = f(\hat{\theta}, u) \tag{1.2}$$

Expression (1.2) implies that, unlike the linear system case which only depends upon the input, the cost function curvature is a function of the location in the parameter space, and it changes depending on the parameter estimate. Therefore, the reduced-order parameter estimation procedure depends upon a leap of faith that the relative importance of the parameters (the relative size of their associated eigenvalues) is more-or-less the same in a reasonably sized region of the parameter space. In the following chapter, this assumption will be shown to be workable in several examples.

2 REDUCED-ORDER PARAMETER ESTIMATION PROCEDURE

This chapter presents the context for a procedure to reduce the number of parameters estimated in a complex, nonlinear process-based plant growth model. The ideas behind the procedure have been used in system identification since the 1960's, but their application to biological growth models is new. Model reduction from a system identification perspective is based upon the premise that a reduced model can closely approximate the dynamics from inputs to outputs of a more complex over-parameterized model. For example, in a successful early example of control theory for a paper pulp mill, rather than developing a complex mechanistic model of the process, input-output data were used to guide model development (Åström 1967) to a low-order linear model with a time delay for control purposes under normal operating conditions. In such approaches, model equations and parameters may have no direct physical interpretation as they do in process-based models.

If one assumes that all states in a complex process-based model need to be measured in order to carry out parameter identification, then the problem of system identification becomes daunting. Although the full system model may have multiple internal states of interest, in the system identification approach presented here any system signal that is measured is considered an output in the reduced model. All inputs are assumed to be measurable and the input data noise free (e.g., solar radiation, temperature). This philosophy, combined with requirement that all inputs and outputs be readily measurable in the field, avoids the overwhelming challenge often noted for parameter identification of complex models with numerous submodels that there are insufficient data for system

identification approaches (Sivakumar 2007; de Reffye et al. 2008; Luo et al. 2009; C. Dupraz, Research Director, INRA, pers. comm., Feb. 10, 2014).

Models are a simplified representation of a system. In this sense, any model is a 'reduced' version of reality. To be clear, the purpose of this procedure is not to improve upon the accuracy of the original process-based system model, but only to systematically determine a reliably identifiable reduced parameter space for estimation, yielding model output behavior that closely matches true model outputs for given classes of inputs.

As noted in Section 1.3.3.1, many process-based models are ill-suited for output prediction due to the large number of parameters and model complexity. A complex model may well include dynamics that have a minor influence on outcomes—model reduction may remove some of these minor modes, while retaining the most important dynamics (Ljung 2010). Along these lines of eliminating minor modes in model reduction, Pearson (2006) writes, "... simplifying assumptions are reasonable if the neglected phenomena have a sufficiently small influence on the behavior of the process that the resulting approximation errors are not too large." The goal of this procedure is to arrive at a reduced-order parameter space that can be estimated entirely from input-output data from the full system (Ljung 1999). Furthermore, the goal is for the reduced model to closely follow the outputs of the original system (with any feasible parameterization) when driven by any input in the input class.

2.1 Model class under consideration

This procedure was developed for application to process-based models such as are common for tree growth (Landsberg and Sands 2011, Weiskittel et al. 2014).

These process-based models are constructed from deterministic descriptions (submodels) characterizing the underlying growth dynamics considered by the modeler to be of importance such as interception and conversion of solar energy, biomass growth, and transpiration. (Buck-Sorlin 2013). Each submodel is a well-studied dynamic process with its own inputs and outputs described by a set of differential or difference equations. Submodel outputs are often inputs to other submodels (Buck-Sorlin 2013). Mathematical representations for the dynamic processes within and between submodels are drawn from the extensive body of work in environmental physics (e.g., Campbell and Norman 2012, Monteith and Unsworth 2013), quantitative agronomy (e.g., Vries et al. 1989, Villalobos and Fereres 2016), and other fields. Submodels may represent model states such as tree leaf area, biomass, and soil water. The models are typically nonlinear in the state variables and in the parameters. As stated by Young (2013) this method of model construction, "... often results in very large simulation models that ... are not fully identifiable from the available data." Due to their tendency toward model complexity, it is assumed that parameter identifiability cannot be resolved using analytic methods.

Basic requirements for I/O system identification include that the model is driven by measurable inputs (e.g., solar radiation, temperature, precipitation) and that all outputs are also measurable (e.g., diameter at breast height [DBH], height) (Ljung 1999). A set of independent variables such as readily obtained system characteristics (such as species, soil type, and location) commonly used in such models (Landsberg and Sands 2011) are considered to be known. Model parameters are typically determined through controlled experimentation at the submodel level and used directly or adjusted based upon expert knowledge

(Young 1983). The physical interpretation of model parameters or the mathematical representation constrain parameter values to a range considered feasible (Hengl et al. 2007, Raue et al. 2011). In other words, parameter value ranges are assumed to be more-or-less known although in order to fit data, these ranges may be exceeded in the parameter estimation phase of model development (Keesman et al. 2011). Additionally, estimates of initial values are needed to run the model, which usually come from biophysical knowledge. The following model reduction procedure locates a lower dimensional parameter space where the parameters are identifiable in practice, but does not question the underlying principles behind the model.

2.2 Establish class of inputs

Model inputs include environmental data such as solar radiation, temperature, rainfall, wind speed, relative humidity, soil nutrients, and so on. Historical weather data are available for locations around the world from various databases (https://www.ncdc.noaa.gov/cdo-web/, http://koeppen-geiger.vu-wien.ac.at/shifts.htm, http://archive.ceda.ac.uk/). However, required data may not be available for the duration (many years) or the sampling interval (e.g., day, month, year) required by the model. A range of stochastic weather data generators has been developed for climate studies that are available for generation of synthetic weather data (Ailliot et al. 2015). Such synthetic data are widely used for crop growth modeling purposes as stated by Ailliot et al. (2015), "Non-linear interactions in process-based models imply that small variations in weather inputs can lead to large output discrepancies, and to counter-intuitive behavior. ... To investigate the influence of weather conditions on such crop

models, it is essential to be able to explore the weather parameter space via simulations."

An additional class of inputs can be called 'management inputs.' These include irrigation, pruning, and soil nutrient amendments. It is assumed that these types of inputs are measurable and that their simulation in the model is straightforward.

2.3 Identifiability

From a system identification viewpoint, it is a desirable characteristic of models that their parameters can uniquely be identified from input-output data, i.e., that their parameters are identifiable. Otherwise the exercise of collecting data would be futile, when model parameters cannot be uniquely determined. Ljung and Glad (1994) state, "It is a fundamental problem of identification to be able—even before the data have been analyzed—to decide if all the free parameters of a model structure can be uniquely recovered from data."

In process-based modeling, non-identifiability may be caused by functional relationships between parameters that are unintentionally created during the model building process due to complex interactions between submodels. Such structural identifiability problems are often not considered during the model-building process due to the costs of model analysis and a reluctance to uncover model problems after a model has been shown to have adequate performance (Crout et al. 2009, Villaverde et al. 2016). Walter (2012) suggests that identifiability is not always required of models, "If one is only interested in reproducing an observed input-output behavior, then it does not matter if there are several models that do the job in exactly the same manner." However, Walter

also affirms that identifiability is required if parameters are expected to have a physical interpretation or if decisions are to be based on parameter values. Many parameters in process-based models are difficult or impossible to measure directly, so their estimation via data is required. Therefore, identifiability is also a convenient model property for reproducibility and comparison of model-building outcomes across different studies.

Identifiability has been a cause for concern in process-based modeling of natural systems for decades. In 1983, Beck voiced concern over parameter identification issues, "Few publications consider this relation in any detail, although it is clear that the uncertainties in a model and its predictions are a function of how the model has been identified and calibrated." Young (1983) also raised an alarm regarding parameter identification in complex process-based models, "The need to choose a model that is efficiently parameterized and compatible with the identifiability of the system (in relation to the available data) is a major requirement of the model-building procedure discussed here: it is clearly foolhardy to attempt the statistical estimation of parameters in a model if the model has excess content (in the form of surplus structure and/or parameters) which cannot be validated against the observed data." Vilela et al. (2009) present an example where a large manifold exists in the parameter space where the cost function is small and note, "Many recent publications have pointed out that multiparametric models tend to have the capacity of accommodating whole ranges of parameter values without much affecting the system dynamics." Furthermore, they determine, "... that there are typically well-defined directions in the parameter space to which the system dynamics is insensitive."

Numerous approaches have been taken to determine global structural identifiability in nonlinear models. Ljung and Glad (1994) propose a differential algebra approach as a means to analyze an arbitrary model structure (with analytic nonlinearities and time-invariant parameters) for global identifiability. The method also addresses issues related to the input requirements for identifiability or "persistent excitation." Audoly et al. (2001) extend differential algebra techniques to an algorithm utilizing computer algebra for testing global identifiability of nonlinear systems such as those commonly found in biological process-based models, including time-varying parameters. Gerdin et al. (2007) suggest a method to determine identifiability of models constructed from submodels, such as are used in the process-based models considered here. The method makes use of the modular model structure to conclude that, "Global identifiability is obtained if and only if each module is identifiable, and the connecting signals can be retrieved from the external signals, without knowledge of the values of the parameters." In system identification terms this means that it is possible to uniquely determine all the model parameters given the measured output signals, or in other words, only one set of parameters can minimize the cost function.

In addition to structural identifiability issues, there is a second, and perhaps more important problem with practical identifiability given the quantity and quality of data reasonably available (Hengl et al. 2007). Practical identifiability problems may arise from several sources

1. Dynamic modes in the model that have little impact on observable outputs.
2. Close functional relationships between parameters that arise from the

model structure but that cannot be detected via analytic methods.

3. Inadequate and/or noisy data.
4. Insufficient richness in the inputs to stimulate all dynamic modes of the model.
5. Failure of model to represent important system dynamics.

This research focuses on the first two of these identifiability issues, which are internal model properties. Items 3–5 are of indirect interest, but are not addressed as a research focus here. A common way to address identifiability of complex models with poor parameter identifiability in practice is to fix the values of certain parameters based on expert knowledge and sensitivity analysis, in the hopes that the cost function hypersurface will have a well-defined minimum for the remaining parameters to be estimated (Young 1983, Sjöberg et al. 2000). To avoid such *ad hoc* approaches, a range of techniques are available to assess identifiability. The Fisher Information Matrix (FIM) is a commonly used tool for assessing parameter uncertainty due to noise in the data (Walter 2012). In a stochastic framework, the FIM gives an *a priori* measure of how much information a model output contains about the parameters that are used to model it.

Arguing that analytic methods for determining identifiability such as those based upon differential algebra "become mathematically intractable with increasing model complexity," Hengl et al. (2007) propose a data based method applicable to realistic experimental conditions. The method utilizes the Alternating Conditional Expectation (ACE)-algorithm (Wang and Murphy 2004) to determine functional relationships between parameters from data generated by simulations. This method reduces the identifiability problem to the problem of finding functionally

related parameters through simulations, then fixing the values of certain parameters based on their functional relationship to other parameters.

The larger the number of parameters, the less likely the correct parameters can be determined, even in perfect identification conditions. Therefore, due to the large number of parameters in many process-based plant growth models, I/O identification of parameters may well be unsuccessful (Cournède et al. 2008, de Reffye et al. 2008, Young and Ratto 2009).

Reduced complexity modeling is often possible when the biophysical descriptions of the system (such as leaf area, number of branches, or root biomass) fall outside the purpose of the model. This may occur when the sole intent of the model is to predict a single output (such as dbh). Giving up the need for knowledge of the internal variables or states can help eliminate expressions that contribute little to the target output, thus leading to a reduced complexity model. The target of the reduced complexity modeling is to identify which parameters can be set to fixed values or which parts of the system can be removed without affecting the target output significantly (e.g., in the previously mentioned paper pulp mill example [Åström 1967]).

2.4 Consistent identifiability

A parametric model is uniquely identifiable if for any parameterization θ^*, the model has identical outputs for any input and time to the model parameterized by $\hat{\theta}$ if and only if $\theta^* = \hat{\theta}$ (Walter 2012). One may begin characterizing the problem of determining identifiability by establishing the model class, input class, and a criterion for comparison (Åström and Eykhoff 1971). When the criterion is a loss function, then the problem of identifiability becomes an optimization

problem that has a unique solution when the system parameters are identifiable. Identifiability has been well-studied for linear systems (Ljung 1999, Walter 2012) and classes of nonlinear systems (e.g., Vilela et al. 2009).

Due to the complex nonlinear nature of many process-based tree growth models (and biological models in general) and frequently noisy data, theoretical identifiability may be much less of a concern than practical identifiability, the ability to uniquely identify parameters in practice. Walter (2012) states that even if parameters are shown to be globally identifiable, it may be more important to know how identifiable they are. As noted above, parameter identifiability depends upon the richness of the inputs (excitation of all system modes) and how well the represents important system dynamics.

For the purposes of this analysis, a model-specific concept is introduced: consistent identifiability. From any point in the reduced parameter space and within the bounds of feasibility defined by the modeler, optimized parameters converge to values within 0.01% of each other. It is impractical to test points covering the entire reduced parameter space. In practice, as demonstrated in Section 2.6, the estimates from two reasonably far apart initializations are compared. Small differences in optimized values from different initializations typically occur due to the algorithm termination tolerances such as for changes in the loss function or parameter values being satisfied from slightly different parameter estimates. For example, if the algorithm is set to terminate if the next step reduces the loss function less than 10^{-5} in magnitude, then a range of points in a small neighborhood of the estimate will also meet the termination criterion.

2.5 Linear basis for procedure

Because of its simplicity, a time invariant moving average or finite impulse response model, can inform our model reduction procedure. The full system we wish to identify is

$$y(n) = \sum_{i=1}^{m} b_i u(n-i) \tag{2.1}$$

with m parameters $b_i \in \mathbb{R}$, $u(n) \in \mathbb{R}$, and time n. The model is chosen as

$$\hat{y}(n) = \sum_{i=1}^{m} \hat{b}_i(n-1) u(n-i) \tag{2.2}$$

with parameter estimates $\hat{b}_i$. Initially, no noise is present in the measurements of output $y(n)$ or the input $u(n)$. The model error is given by

$$r(n) = y(n) - \hat{y}(n) \tag{2.3}$$

The cost function is defined as the squared error

$$\begin{aligned} V(n) &= \frac{1}{2} r^2(n) \\ &= \frac{1}{2}\left(y(n) - \hat{y}(n)\right)^2 \\ &= \frac{1}{2}\left(y(n) - \sum_{i=1}^{m} \hat{b}_i(n-1) u(n-i) \right)^2 \end{aligned} \tag{2.4}$$

For a given set of measurements of $y(n)$ and $u(n)$, (2.4) defines a quadratic surface in the parameter estimates $\hat{b}_i$. Beginning with an *a priori* parameter estimate (our best guess), one would like to descend this error surface towards the actual (or true) parameters until the $\hat{b}_i$ are very close to the b_i. To this end, the upward gradient of the cost function for the $\hat{b}_i^{\text{th}}$ parameter is

$$\frac{\partial V(n)}{\partial \hat{b}_i(n-1)} = -u(n-i)r(n) \tag{2.5}$$

This leads to the well-known least mean square (LMS) recursive algorithm to descend the gradient of the cost function (2.4) (Widrow and Stearns 1985):

$$\begin{aligned} \hat{b}_i(n) &= \hat{b}_i(n-1) - \mu \frac{\partial V(n)}{\partial \hat{b}_i(n-1)} \\ &= \hat{b}_i(n-1) + \mu u(n-i)r(n) \end{aligned} \tag{2.6}$$

where the update term on the RHS of (2.6) is multiplied by a small positive term μ that influences the speed and stability of the algorithm. The LMS algorithm owes its popularity to its computational simplicity, straightforward interpretation, and robustness (Mendel 1973, Widrow and Stearns 1985, Slock 1993).

If the Hessian matrix of second partial derivatives of a function with respect to the independent variables (parameters) is positive definite, then the function has only one extreme point (Bellman and Åström 1970). A positive definite m x m matrix H is positive definite if for all m x 1 vectors $a \neq 0$

$$a^T H a > 0 \tag{2.7}$$

where a^T is the transpose of a. An equivalent property of positive definite matrices is that all its eigenvalues are positive.

The Hessian matrix of the second partial derivatives of a function f is of the form

$$H(f)=\begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_m} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_m \partial x_1} & \frac{\partial^2 f}{\partial x_m \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_m^2} \end{bmatrix} \tag{2.8}$$

where in our case f is the cost function of interest $V(n)$ and x_k are the m parameter estimates. For our case the i,j^{th} term of (2.8) is

$$\begin{aligned} \frac{\partial^2 V(n)}{\partial \hat{b}_i(n-1)\partial \hat{b}_j(n-1)} &= \frac{\partial}{\partial \hat{b}_j(n-1)}\left(\frac{\partial V(n)}{\partial \hat{b}_i(n-1)}\right) \\ &= \frac{\partial}{\partial \hat{b}_j(n-1)}\left(-u(n-i)r(n)\right) \\ &= -\frac{\partial u(n-i)}{\partial \hat{b}_j(n-1)}r(n) - u(n-i)\frac{\partial r(n)}{\partial \hat{b}_j(n-1)} \end{aligned} \tag{2.9}$$

Since the input $u(n)$ is independent of the parameter estimates, (2.9) further reduces to

$$\begin{aligned} \frac{\partial^2 V(n)}{\partial \hat{b}_i(n-1)\partial \hat{b}_j(n-1)} &= -u(n-i)\frac{\partial r(n)}{\partial \hat{b}_j(n-1)} \\ &= u(n-i)u(n-j) \end{aligned} \tag{2.10}$$

The Hessian in the case of an m-order FIR model then becomes

$$H(V(n))=\begin{bmatrix} u^2(n-1) & u(n-1)u(n-2) & \cdots & u(n-1)u(n-m) \\ u(n-2)u(n-1) & u^2(n-2) & \cdots & u(n-2)u(n-m) \\ \vdots & \vdots & \ddots & \vdots \\ u(n-m)u(n-1) & u(n-m)u(n-2) & \cdots & u^2(n-m) \end{bmatrix} \tag{2.11}$$

Note that this Hessian is only dependent upon the input and not the parameter estimates. In other words, for this model the Hessian is the same throughout the parameter space and only a property of the regressors (Sjöberg et al. 1995). Input

excitation can determine parameter identifiability and has been treated throughout the system identification literature (Åström and Eykhoff 1971, Godfrey and Brown 1979, Ljung 1999, Keesman 2011).

The requirement for there to be a unique minimum for the quadratic cost function is that the Hessian be positive definite. However, the expression (2.11) is an outer product of a vector, which always has rank one for non-trivial inputs and m–1 zero eigenvalues. To address this deficit in rank, one clearly must consider the cost $V(n)$ at more than one point in time. Therefore, consider the estimation scheme of (2.6) averaged over time. The concatenated set of m versions of (2.6) is

$$\hat{\theta}(n) = \hat{\theta}(n-1) - \mu X(n-1) r(n) \tag{2.12}$$

where the vector of parameter estimates is

$$\hat{\theta}(n) = \begin{bmatrix} \hat{b}_1(n) & \hat{b}_2(n) & \cdots & \hat{b}_m(n) \end{bmatrix}^T \tag{2.13}$$

and the vector of past inputs, the regressor, is defined as

$$X(n-1) = \begin{bmatrix} u(n-1) & u(n-2) & \cdots & u(n-m) \end{bmatrix}^T \tag{2.14}$$

Since the update step size μ and the error $r(n)$ are scalars, one can see that at each iteration of (2.12) the parameter estimate moves in the direction of the regressor $X(n–1)$. This suggests intuitively that the regressor directions are important to convergence properties. To get further along in examining LMS convergence, rephrase the optimization problem as the minimization of the average cost over N steps

$$
\begin{aligned}
V_N(n) &= \frac{1}{N}\sum_{\tau=n-N+1}^{n}\frac{1}{2}r^2(\tau) \\
&= \frac{1}{N}\sum_{\tau=n-N+1}^{n}\frac{1}{2}\left(y(\tau)-\hat{y}(\tau)\right)^2 \\
&= \frac{1}{N}\sum_{\tau=n-N+1}^{n}\frac{1}{2}\left(y(\tau)-\sum_{i=1}^{m}\hat{b}_i(\tau-i)u(\tau-i)\right)^2
\end{aligned}
\tag{2.15}
$$

Following the steps of (2.5) and (2.10) above to derive the Hessian of the average cost, the first partial derivatives of $V_N(n)$ with respect to $\hat{b}_i(n-1)$ are

$$
\frac{\partial V_N(n)}{\partial \hat{b}_i(n-1)} = \frac{1}{N}\sum_{\tau=n-N+1}^{n} -u(\tau-i)r(\tau) \tag{2.16}
$$

and the second partial derivatives (the terms of the Hessian) are

$$
\frac{\partial^2 V_N(n)}{\partial \hat{b}_i(n-1)\partial \hat{b}_j(n-1)} = \frac{1}{N}\sum_{\tau=n-N+1}^{n} u(\tau-i)u(\tau-j) \tag{2.17}
$$

Now consider an example 2-parameter FIR filter

$$
y(n) = b_1 u(n-1) + b_2 u(n-2) \tag{2.18}
$$

with parameters $b_1 = b_2 = 1$. With this parameter setting, (2.18) is a simple filter that passes low frequency inputs and attenuates high frequencies. One can see this intuitively, as low frequencies tend to have similar sample-to-sample values that are additive, while high frequencies have different sample-to-sample values with a relatively lower average value.

To illustrate how the input influences the curvature of the cost function, consider an input composed of two sinusoids

$$
u(n) = A_1 \sin\left(\frac{2\pi\omega_1 n}{N}\right) + A_2 \sin\left(\frac{2\pi\omega_2 n}{N}\right) \tag{2.19}
$$

of amplitudes A_1 and A_2, frequencies ω_1 and ω_2, and step size n and simulation length N. We then compare LMS parameter estimates for two trial inputs that

combine low and high frequency sinusoids (Table 2.1). For these simulations a number of initial parameter estimates are selected, as given in Table 2.2.

Table 2.1 Coefficients for trial inputs (2.19).

Trial input	A_1	A_2	ω_1	ω_2	N
u_1	5	1	2	107.1	1000
u_2	1	5	2	107.1	1000

Table 2.2 Initial parameter estimates for 2-parameter FIR model.

A priori estimate	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\theta}_5$	$\hat{\theta}_6$	$\hat{\theta}_7$
$\hat{b}_1$	0.5	0.5	0.7	1	1.2	1.5	1.5
$\hat{b}_2$	0.5	1	1.5	0.5	1.5	0.7	1.2

The results of implementing the LMS (μ=0.005, number of steps = 500) for the two inputs are given in Figure 2.1, which shows the trajectories from all seven initial parameter estimates superimposed on the contours (equi-level curves) of the cost function. For Input 2, the cost function surface can be described as an elliptical hyperbolic 'bowl' with a unique minimum at the correct parameter location and LMS has no trouble descending to the correct parameter position. However, for Input 1, there exists a low narrow 'valley' at the bottom of the bowl along the b_1=$-b_2$ axis, where the cost is low and the estimation progress is slow. From all initial parameter estimates the cost is reduced to a small number within 60 steps even though the parameter estimates are far from their actual values. In more complex nonlinear models, such slow modes in the cost function surface for a given class of inputs. As will be demonstrated for a 12-parameter model in Section 3.4.1, stalling of the parameter estimates in a slow mode can be avoided by reducing the dimension of the parameter estimate space by fixing the value of certain parameters and estimating others.

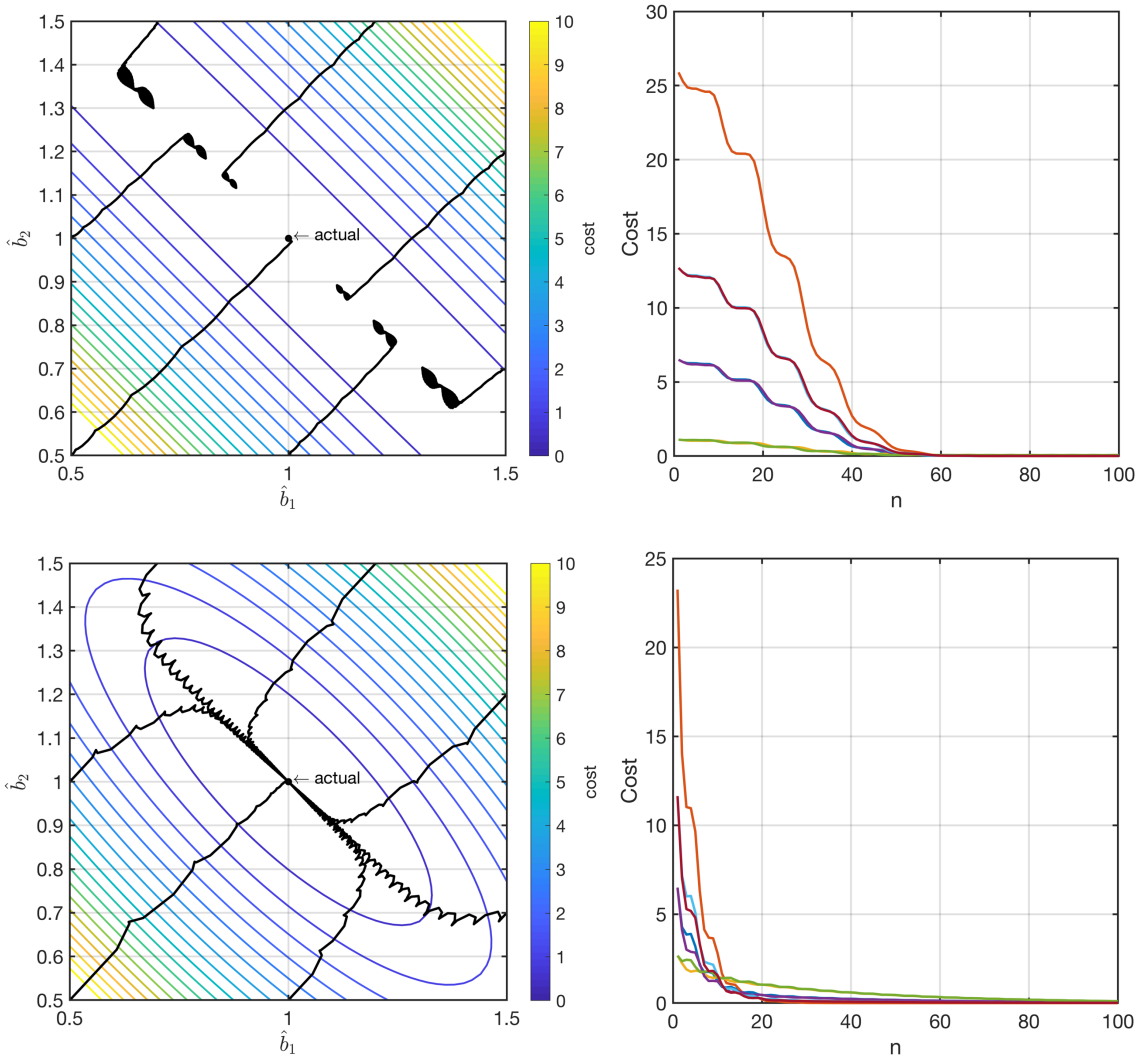

Figure 2.1 Contours and LMS parameter estimate trajectories (left) and cost function values (right) for the seven initial parameter estimates and two inputs (Input 1, top; Input 2, bottom). The length of the simulation is 500 steps (μ=0.005), although only the first 100 steps are shown in the cost function values (right), as the values quickly approach very small numbers.

To gain insight into the findings of Figure 2.1, calculate the Hessian for Inputs 1 and 2. Because our inputs are the sum of sinusoids, the average of the terms in (2.11) could be determined analytically via trigonometric identities. However, since the goal is to apply well established concepts from linear systems to nonlinear systems, here the average Hessian is computed numerically. Combining

(2.11) and (2.17) into an expression for the average Hessian of a 2-parameter FIR model gives

$$\bar{H}(V_N(n)) = \begin{pmatrix} \frac{1}{N}\sum_{\tau=n-N+1}^{n} u^2(\tau-1) & \frac{1}{N}\sum_{\tau=n-N+1}^{n} u(\tau-1)u(\tau-2) \\ \frac{1}{N}\sum_{\tau=n-N+1}^{n} u(\tau-2)u(\tau-1) & \frac{1}{N}\sum_{\tau=n-N+1}^{n} u^2(\tau-2) \end{pmatrix} \quad (2.20)$$

$$= \frac{1}{N}\sum_{\tau=n-N+1}^{n} \begin{pmatrix} u^2(\tau-1) & u(\tau-1)u(\tau-2) \\ u(\tau-2)u(\tau-1) & u^2(\tau-2) \end{pmatrix}$$

From this expression, one can directly calculate the average Hessian over N steps of the input (Table 2.3). Averaging techniques as presented in Johnson (1988) show that the axes of the ellipses of the cost function contours shown in Figure 2.1 are oriented along the eigenvectors of the average Hessian and that the shape of the ellipses is determined by the relative magnitude of the eigenvalues. This observation is key to interpreting the curvature of the cost function based upon the Hessian. The eigenstructure is determined by the solution to

$$\mathrm{H}^* \mathrm{X} = \mathrm{X}^* D \quad (2.21)$$

where D is the m x m diagonal matrix of eigenvalues λ_i and X is a m x m matrix whose columns are the corresponding right eigenvectors, such that $\mathrm{H}^* \mathrm{X} = \mathrm{X}^* D$. In our calculations using MATLAB's eig function, the eigenvectors are always normalized to unit length. The eigenstructure of our 2-parameter model for the two inputs is given in Table 2.3.

Table 2.3 Hessian, eigenvalues, and eigenvectors for Inputs 1 and 2 (N=1000).

	Hessian Input 1		**Hessian Input 2**	
	6.0643	6.0064	5.5990	4.4319
	6.0064	6.0419	4.4319	5.5835
	Eigenvalue		**Eigenvalue**	
	λ_1	λ_2	λ_1	λ_2
	1.21e+01	4.67e-02	1.00e+01	1.16e+00
	Eigenvector		**Eigenvector**	
Axis	X_1	X_2	X_1	X_2
b_1	-0.7078	0.7064	-0.7077	0.7065
b_2	-0.7064	-0.7078	-0.7065	-0.7077

Next consider what happens to the cost function surface if independent, normally distributed noise (ε) is introduced into the model output

$$y(n) = b_1 u(n-1) + b_2 u(n-2) + \varepsilon(n) \tag{2.22}$$

Because noise is not a function of the parameters, its derivatives with respect to the parameters do not appear in the Hessian, i.e., the noise does not affect the curvature of the loss function. Instead, noise has the effect of lifting the bowl so that the average minimum cost is greater than 0. A simulation to demonstrate this effect is shown in Figure 2.2 with $b_1 = b_2 = 1$ and ε(n) normally distributed with amplitude 0.5. Note that the cost does not reach 0, which is reflected in the parameter estimate jittering due to the noise.

Finally, for the linear case, consider what happens if there is error in the model (and no noise in the output). We begin with an FIR model that is the same as the previous one, but with a third term that has a negligible effect on the output

$$y(n) = b_1 u(n-1) + b_2 u(n-2) + b_3 u(n-3) \tag{2.23}$$

To assure that the new term is relatively insignificant, again set $b_1 = b_2 = 1$ and $b_3 = 0.1$.

The output of the 3-parameter (2.23) is considered to be the true system, while our approximation is the 2-parameter model

$$\hat{y}(n) = \hat{b}_1 u(n-1) + \hat{b}_2 u(n-2) \tag{2.24}$$

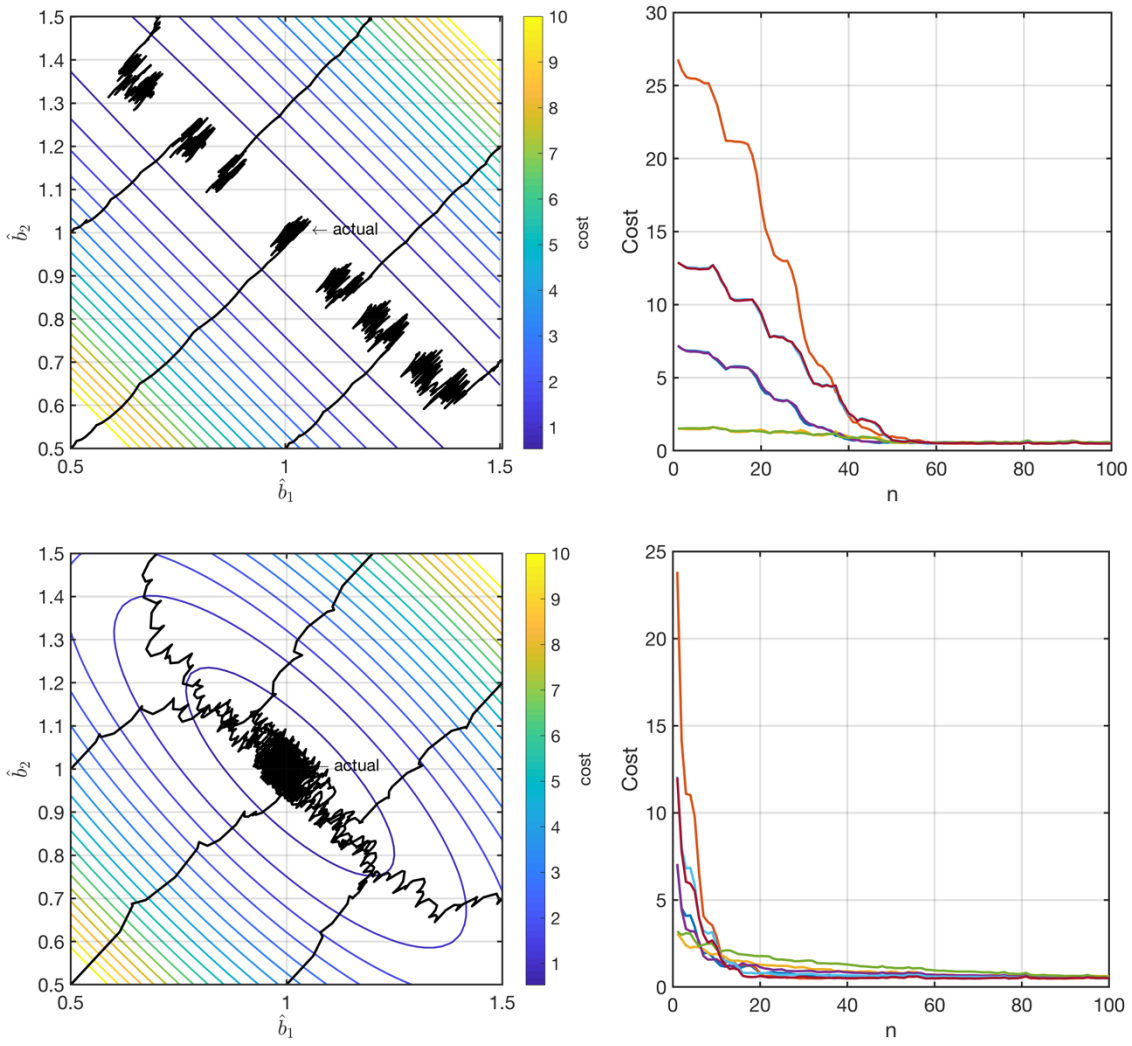

Figure 2.2 Contours and LMS parameter estimate trajectories (left) and cost function values (right) for the seven initial parameter estimates and two inputs (Input 1, top; Input 2, bottom) with noise in the output. The length of the simulation is 500 steps (μ=0.005), with the first 100 steps shown in the cost function values (right).

One can see from Figure 2.3 that model error shifts the location of the cost surface so that the minimum no longer lies at the true parameter location, but

some distance away. This demonstrates that when there is model error, the cost function minimum will move to a location away from the true parameters where the estimated parameters compensate as much as they can for deficiencies in the model structure. Note that the lack of excitation by Input 1 still leads to slow convergence to the cost minimum. In the case of Input 2, the estimate converges to the minimum, but the cost does not reach 0 due to the model error.

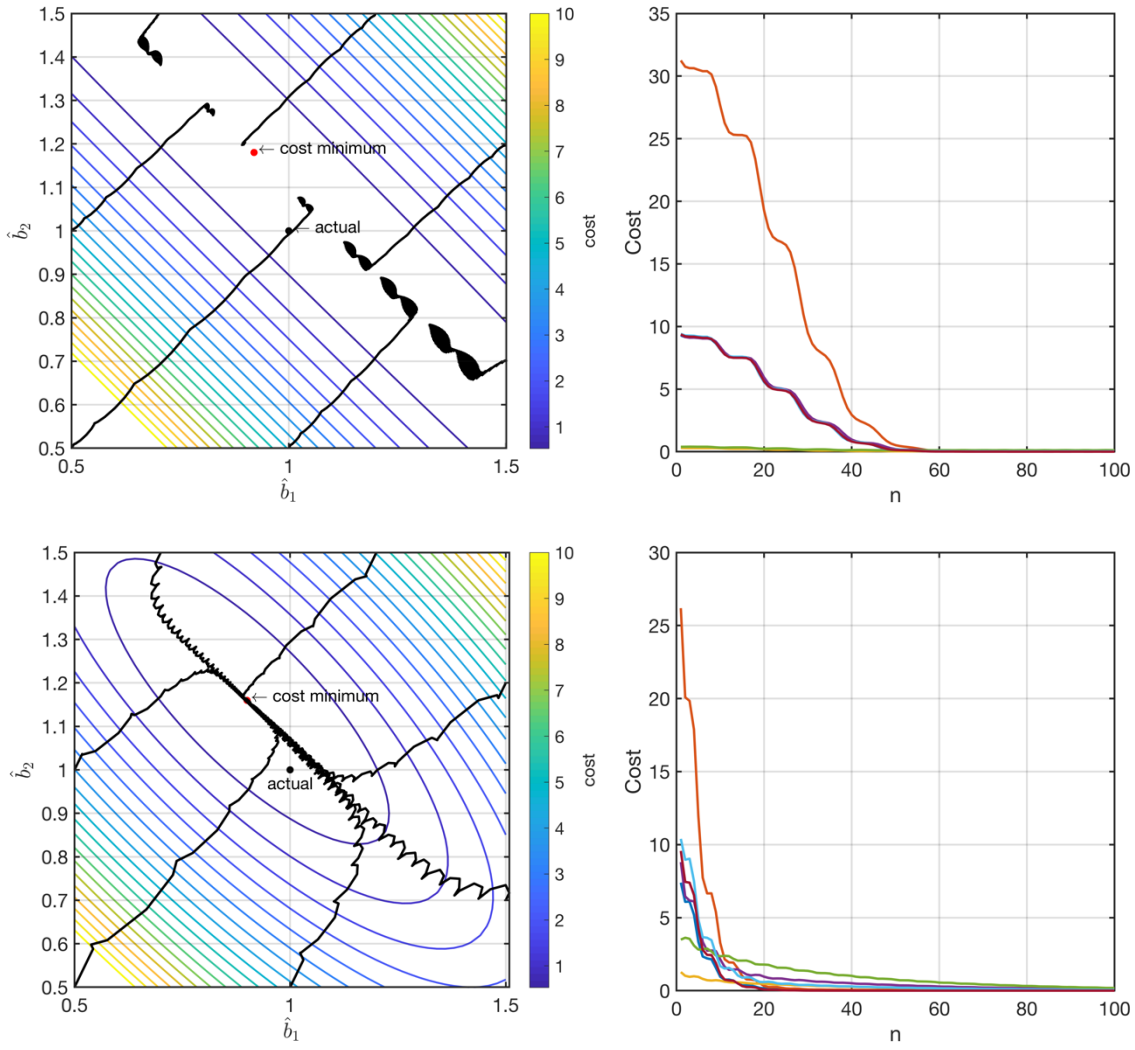

Figure 2.3 Contours and LMS parameter estimate trajectories (left) and cost function values (right) for the seven initial parameter estimates and two inputs (Input 1, top; Input 2, bottom) with model error. The length of the simulation is 500 steps (μ=0.005), with 100 steps shown in the cost function values (right).

2.6 Nonlinear example

Now consider a commonly used growth model, the logistic or Verhulst equation, to see how concepts explored in the linear case carry over to a nonlinear case. Consider the a 3-parameter logistic equation

$$y(t) = \frac{b_1}{1 + b_2 \, \mathrm{e}^{-b_3 t}} \tag{2.25}$$

with time t and parameters b_1, b_2, and b_3 positive real in our example. There is no exogenous input to this equation, although one might consider time (t) to be the driving input. The parameters are set at a range of magnitudes: b_1=1000, b_2=10, and b_3=0.004 to observe how this effects LMS analysis. This parameter setting gives the trajectory of $y(t)$ for $t = 0 - 1000$ shown in Figure 2.4. We also pick a set of initial parameter values called '*a priori*' estimates or a best guess of the parameter values based on prior experience or expert knowledge. The actual and a priori parameters are given in Table 2.4.

Table 2.4 Parameters and actual and *a priori* estimated values used in the logistic equation example.

Parameter	**Actual value**	**Estimated parameters**	***A priori* value**
b_1	1000	$\hat{b}_1$	950
b_2	10	$\hat{b}_2$	9.5
b_3	0.004	$\hat{b}_3$	0.0036

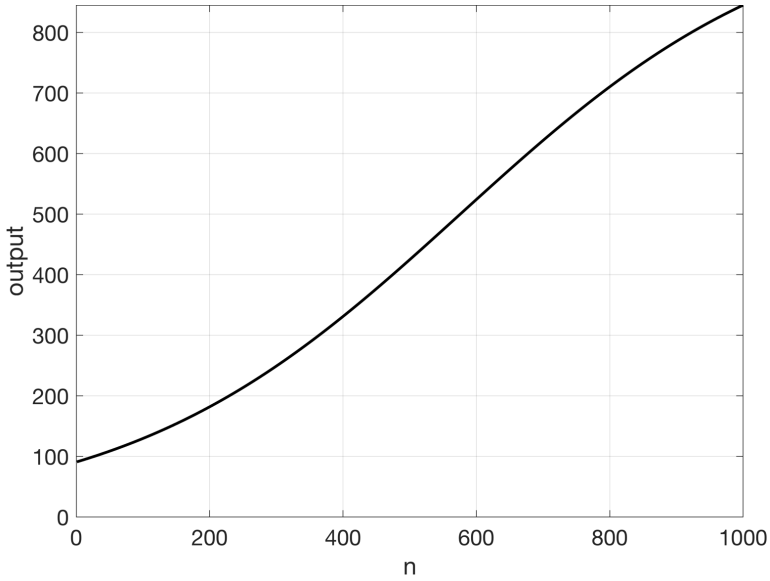

Figure 2.4 Output of logistic equation (2.25) with b_1=1000, b_2=10, and b_3=0.004.

The model is of identical form to (2.25) with *a priori* parameter estimates $\hat{b}_1$, $\hat{b}_2$, and $\hat{b}_3$

$$\hat{y}(t) = \frac{\hat{b}_1}{1 + \hat{b}_2 \, \mathrm{e}^{-\hat{b}_3 t}} \tag{2.26}$$

The model error at any time t is given by

$$\begin{aligned} r(t) &= y(t) - \hat{y}(t) \\ &= y(t) - \frac{\hat{b}_1}{1 + \hat{b}_2 \, \mathrm{e}^{-\hat{b}_3 t}} \end{aligned} \tag{2.27}$$

leading to an expression for the instantaneous cost function

$$\begin{aligned} V(t) &= \frac{1}{2} r^2(t) \\ &= \frac{1}{2}\left(y(t) - \frac{\hat{b}_1}{1 + \hat{b}_2 \, \mathrm{e}^{-\hat{b}_3 t}} \right)^2 \end{aligned} \tag{2.28}$$

As demonstrated in the linear case, the average cost over a period of time is of interest, rather than the instantaneous cost. In this case, take the average over discrete steps $t = 1, 2, \ldots, N$, with time index i expressed as

$$
\begin{aligned}
V_N(n) &= \frac{1}{N}\sum_{i=n-N+1}^{n}\frac{1}{2}r^2(i) \\
&= \frac{1}{N}\sum_{i=n-N+1}^{n}\frac{1}{2}\left(y(i)-\frac{\hat{b}_1}{1+\hat{b}_2\,\mathrm{e}^{-\hat{b}_3 i}}\right)^2
\end{aligned}
\tag{2.29}
$$

2.6.1 Hessian formulation

The Hessian matrix of the second partial derivatives of a function f is of the form

$$
H(f)=\begin{bmatrix}
\dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1\,\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1\,\partial x_m} \\
\dfrac{\partial^2 f}{\partial x_2\,\partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2\,\partial x_m} \\
\vdots & \vdots & \ddots & \vdots \\
\dfrac{\partial^2 f}{\partial x_m\,\partial x_1} & \dfrac{\partial^2 f}{\partial x_m\,\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_m^2}
\end{bmatrix}
\tag{2.30}
$$

where in our case f is the cost function of interest is (2.29) and x_k are the m parameters, for which the general form of the average Hessian was given in (2.20) , leading to a general form of the average Hessian for N discrete time points

$$
H(V_N)=\frac{1}{2N}\begin{bmatrix}
\dfrac{\partial^2}{\partial x_1^2}\displaystyle\sum_{i=1}^{N}r^2(i) & \dfrac{\partial^2}{\partial x_1\,\partial x_2}\displaystyle\sum_{i=1}^{N}r^2(i) & \cdots & \dfrac{\partial^2}{\partial x_1\,\partial x_m}\displaystyle\sum_{i=1}^{N}r^2(i) \\
\dfrac{\partial^2}{\partial x_2\,\partial x_1}\displaystyle\sum_{i=1}^{N}r^2(i) & \dfrac{\partial^2}{\partial x_2^2}\displaystyle\sum_{i=1}^{N}r^2(i) & \cdots & \dfrac{\partial^2}{\partial x_2\,\partial x_m}\displaystyle\sum_{i=1}^{N}r^2(i) \\
\vdots & \vdots & \ddots & \vdots \\
\dfrac{\partial^2}{\partial x_m\,\partial x_1}\displaystyle\sum_{i=1}^{N}r^2(i) & \dfrac{\partial^2}{\partial x_m\,\partial x_2}\displaystyle\sum_{i=1}^{N}r^2(i) & \cdots & \dfrac{\partial^2}{\partial x_m^2}\displaystyle\sum_{i=1}^{N}r^2(i)
\end{bmatrix}
\tag{2.31}
$$

where for less cluttered notation and without loss of generality the time indices are changed to calculate the Hessian for any N consecutive points in the error sequence.

By the sum rule in differentiation, the Hessian of the sum of the squared residuals with respect to the parameters is equal to the sum of the Hessians of the squared residuals, i.e.,

$$H\left[\frac{1}{2}\sum_{i=1}^{N} r^2(i)\right] = \frac{1}{2}\sum_{i=1}^{N} H\left[r^2(i)\right] \tag{2.32}$$

from which (2.31) leads to

$$H(V_N) = \frac{1}{N}\sum_{i=1}^{N}\begin{bmatrix} \frac{1}{2}\frac{\partial^2 r^2(i)}{\partial x_1^2} & \frac{1}{2}\frac{\partial^2 r^2(i)}{\partial x_1 \partial x_2} & \cdots & \frac{1}{2}\frac{\partial^2 r^2(i)}{\partial x_1 \partial x_m} \\ \frac{1}{2}\frac{\partial^2 r^2(i)}{\partial x_2 \partial x_1} & \frac{1}{2}\frac{\partial^2 r^2(i)}{\partial x_2^2} & \cdots & \frac{1}{2}\frac{\partial^2 r^2(i)}{\partial x_2 \partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}\frac{\partial^2 r^2(i)}{\partial x_m \partial x_1} & \frac{1}{2}\frac{\partial^2 r^2(i)}{\partial x_m \partial x_2} & \cdots & \frac{1}{2}\frac{\partial^2 r^2(i)}{\partial x_m^2} \end{bmatrix} \tag{2.33}$$

The Hessian (2.33) can be interpreted as an integration of the second partial derivatives of the residuals over N steps, containing information about the average local curvature of the cost function.

The jth row and kth column term of the Hessian in (2.33) is determined by

$$H_{j,k}(V_N) = \frac{1}{2}\frac{\partial^2 r^2(i)}{\partial x_j \partial x_k} \tag{2.34}$$

The first derivative of $\frac{1}{2}$ the squared residual with respect to the jth parameter is

$$\frac{1}{2}\frac{\partial r^2(i)}{\partial x_j} = \frac{\partial r(i)}{\partial x_j} r(i) \tag{2.35}$$

Which by the product rule leads to the second derivative with respect to the kth parameter

$$\frac{1}{2}\frac{\partial^2 r^2(i)}{\partial x_j \partial x_k} = \frac{\partial r(i)}{\partial x_j}\frac{\partial r(i)}{\partial x_k} + \frac{\partial^2 r(i)}{\partial x_j \partial x_k} r(i) \tag{2.36}$$

which is a convenient expression for the Hessian elements based on first and second partial derivatives of the residual $r(i)$. Assuming $r(i)$ is a continuous twice-differentiable function with respect to the parameters of interest, symmetry of the second derivatives holds, giving

$$H_{j,k}(V_N) = H_{k,j}(V_N) = \frac{\partial r(i)}{\partial x_j}\frac{\partial r(i)}{\partial x_k} + \frac{\partial^2 r(i)}{\partial x_j \partial x_k} r(i) \tag{2.37}$$

which means that elements of the Hessian can be broken down into smaller components: the first and second order partial derivatives of the residual, rather than second order partial derivatives of the squared residual (which are more cumbersome to display here). For the logistic expression (2.25), these components are the first partial derivatives of $r(i)$

$$\begin{aligned}
\frac{\partial r(i)}{\partial \hat{b}_1} &= \frac{1}{1 + \hat{b}_2 \, \mathrm{e}^{-\hat{b}_3 i}} \\
\frac{\partial r(i)}{\partial \hat{b}_2} &= \frac{\hat{b}_1 \, \mathrm{e}^{-\hat{b}_3 i}}{\hat{b}_2 \left(1 + \mathrm{e}^{-\hat{b}_3 i}\right)^2} \\
\frac{\partial r(i)}{\partial \hat{b}_3} &= -\frac{2i\hat{b}_1 \hat{b}_2 \, \mathrm{e}^{-\hat{b}_3 i}}{\left(1 + \hat{b}_2 \, \mathrm{e}^{-\hat{b}_3 i}\right)^2}
\end{aligned} \tag{2.38}$$

and the second partial derivatives of $r(i)$

$$\frac{\partial^2 r(i)}{\partial \hat{b}_1^2} = 0$$

$$\frac{\partial^2 r(i)}{\partial \hat{b}_2^2} = -\frac{2\,\mathrm{e}^{-2\hat{b}_3 i}\,\hat{b}_1}{\left(1+\hat{b}_2\,\mathrm{e}^{-\hat{b}_3 i}\right)^3}$$

$$\frac{\partial^2 r(i)}{\partial \hat{b}_3^2} = -\hat{b}_1\left(\frac{2\,\mathrm{e}^{-2\hat{b}_3 i}\,i^2\hat{b}_2^2}{\left(1+\hat{b}_2\,\mathrm{e}^{-\hat{b}_3 i}\right)^3} - \frac{\mathrm{e}^{-\hat{b}_3 i}\,i^2\hat{b}_2}{\left(1+\hat{b}_2\,\mathrm{e}^{-\hat{b}_3 i}\right)^2}\right)$$

$$\frac{\partial^2 r(i)}{\partial \hat{b}_1\,\partial \hat{b}_2} = \frac{\mathrm{e}^{-\hat{b}_3 i}}{\left(1+\hat{b}_2\,\mathrm{e}^{-\hat{b}_3 i}\right)^2}$$

$$\frac{\partial^2 r(i)}{\partial \hat{b}_2\,\partial \hat{b}_3} = -\frac{\mathrm{e}^{-\hat{b}_3 i}\,i\hat{b}_2}{\left(1+\hat{b}_2\,\mathrm{e}^{-\hat{b}_3 i}\right)^2}$$

$$\frac{\partial^2 r(i)}{\partial \hat{b}_1\,\partial \hat{b}_3} = \frac{2\,\mathrm{e}^{-2\hat{b}_3 i}\,i\hat{b}_1\hat{b}_2}{\left(1+\hat{b}_2\,\mathrm{e}^{-\hat{b}_3 i}\right)^3} - \frac{\mathrm{e}^{-\hat{b}_3 i}\,i\hat{b}_1}{\left(1+\mathrm{e}^{-i\hat{b}_3}\,\hat{b}_2\right)^2} \tag{2.39}$$

Now substitute the partial derivatives of (2.38) and (2.39) into each term (2.37) of the Hessian matrix (2.33) in order to get an analytic expression that can be used to numerically calculate the Hessian for any choice of parameters in the feasible parameter space (positive real values in this example).

Using the selected actual parameters from Table 2.4 one can calculate the Hessian at the actual parameter location (estimated parameters = actual parameters). The Hessian and its eigenstructure at the actual parameters are given in Table 2.5. One can see that the Hessian is ill-conditioned with a condition number of 7e–13, which can present numerical problems in the optimization and stalling of the parameter estimate. One can see from the contour lines of the loss function in the b_1–b_3 space close to the correct parameters that the dominant curvature is along the b_3 direction (Figure 2.5), which corresponds to the dominant eigenvalue of 1.35e+10.

Table 2.5 Hessian, eigenvalues, and eigenvectors for logistic equation example at actual parameter location without parameter scaling.

	Hessian (N=1000)		
	2.54e–01	–8.83e+00	5.76e+04
	–8.83e+00	3.80e+02	–2.11e+06
	5.76e+04	–2.10e+06	1.36e+10
	Eigenvalue		
	λ_1	λ_2	λ_3
	1.36e+10	5.36e+01	9.62e–03
	Eigenvector		
Axis	X_3	X_2	X_1
b_1	0.0000	–0.0020	–1.0000
b_2	–0.0002	–1.0000	0.0020
b_3	1.0000	–0.0002	0.0000

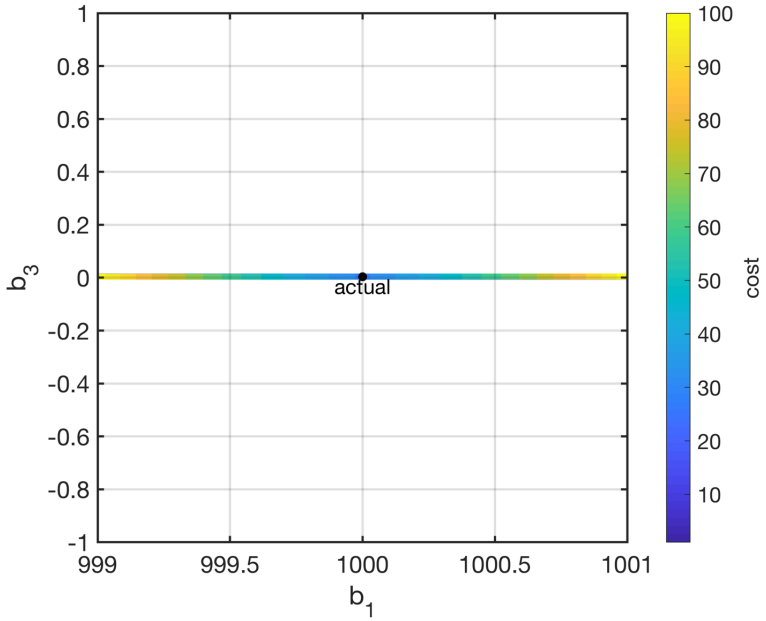

Figure 2.5 Contour lines for logistic equation b_1–b_3 space without parameter scaling.

The first suspect for the source of this ill-conditioning is parameter scaling. Poor choice of parameter scaling can cause an ill-conditioned Hessian (Thacker 1989), as will be shown to be the case here. As shown in a parameter estimation example for the logistic equation by Nash (2010), scaling can greatly impact the eigenstructure of the Hessian. Parameter scaling is commonly used to reduce

numerical problems caused by very differently scaled parameters in the optimization algorithm. In regards to such 'badly scaled' problems using the Trust Region Reflective algorithm, the optimization method used in this study as implemented via MATLAB's `lsqnonlin` option, Conn et al. (2000) state, "It is ... of paramount importance to rescale the problem's variables to make their typical values of comparable magnitude, if at all possible." Others recommend that parameter units are selected such that all parameter values fall at roughly the same magnitude, preferably around 1 (Thacker 1989, Dennis and Schnabel 1996). However, if normalizing parameter values via their units is infeasible, the parameters can be transformed for analysis by a linear transformation in the parameter space. Such scaling in the parameter space improves conditioning of the parameter estimation problem (Coleman and Li 1996).

For rescaling we use a simple substitution in our model

$$\hat{b}_i = s_i \hat{b}'_i \tag{2.40}$$

where a scaling factor s_i multiplies the new parameter $\hat{b}'_i$. The scaling factor s_i is set equal to the actual parameter value (or the *a priori* parameter estimate), while the new parameter $\hat{b}'_i$ has the value of 1. Our scaled model becomes

$$\hat{y}'(t) = \frac{s_1 \hat{b}'_1}{1 + s_2 \hat{b}'_2 \, \mathrm{e}^{-s_3 \hat{b}'_3 t}} \tag{2.41}$$

Note that this rescaling does not change the model output, therefore, to avoid confusion with actual versus scaled parameter values we will plot in the original parameter space, which is by definition the same as the new parameter space with axes scaled by the s_i values. The residual with scaling is

$$
\begin{aligned}
r'(t) &= y(t) - \hat{y}'(t) \\
&= y(t) - \frac{s_1 \hat{b}_1'}{1 + s_2 \hat{b}_2' \, \mathrm{e}^{-s_3 \hat{b}_3' t}}
\end{aligned}
\tag{2.42}
$$

Once again, the scaling of the residual does not change its value. The scaling will change the expressions for the Hessian, as we are now taking partial derivatives with respect to the new scaled parameters $\hat{b}_i'$. The new first order partial derivatives of the residual (2.42) become

$$
\begin{aligned}
\frac{\partial r'(i)}{\partial \hat{b}_1'} &= -\frac{s_1}{1 + s_2 \hat{b}_2' \, \mathrm{e}^{-s_3 \hat{b}_3' i}} \\
\frac{\partial r'(i)}{\partial \hat{b}_2'} &= \frac{s_1 s_2 \hat{b}' \, \mathrm{e}^{-s_3 \hat{b}_3' i}}{\left(1 + s_2 \hat{b}_2' \, \mathrm{e}^{-s_3 \hat{b}_3' i}\right)^2} \\
\frac{\partial r'(i)}{\partial \hat{b}_3'} &= -\frac{s_1 s_2 s_3 \hat{b}_1' \hat{b}_2' \, \mathrm{e}^{-s_3 \hat{b}_3' i} \, i}{\left(1 + s_2 \hat{b}_2' \, \mathrm{e}^{-s_3 \hat{b}_3' i}\right)^2}
\end{aligned}
\tag{2.43}
$$

and the second partial derivations are with scaling

$$
\begin{aligned}
\frac{\partial^2 r'(i)}{\partial \hat{b}_1'^2} &= 0 \\
\frac{\partial^2 r'(i)}{\partial \hat{b}_2'^2} &= \frac{s_1 s_2 \hat{b}_1' \mathrm{e}^{-s_3 \hat{b}_3' i}}{\left(1 + s_2 \hat{b}_2' \mathrm{e}^{-s_3 \hat{b}_3' i}\right)^2} \\
\frac{\partial^2 r'(i)}{\partial \hat{b}_3'^2} &= -s_1 \hat{b}_1' i^2 \left(\frac{2 s_2^2 s_3^2 \hat{b}_2'^2 \mathrm{e}^{-2 s_3 \hat{b}_3' i}}{\left(1 + s_2 \hat{b}_2' \mathrm{e}^{-s_3 \hat{b}_3' i}\right)^3} - \frac{s_2 s_3^2 \hat{b}_2' \mathrm{e}^{-i s_3 \hat{b}_3'}}{\left(1 + s_2 \hat{b}_2' \mathrm{e}^{-s_3 \hat{b}_3' i}\right)^2} \right) \\
\frac{\partial^2 r'(i)}{\partial \hat{b}_1' \partial \hat{b}_2'} &= \frac{s_1 s_2 \mathrm{e}^{-s_3 \hat{b}_3' i}}{\left(1 + s_2 \hat{b}_2' \mathrm{e}^{-s_3 \hat{b}_3' i}\right)^2} \\
\frac{\partial^2 r'(i)}{\partial \hat{b}_2' \partial \hat{b}_3'} &= -\frac{s_1 s_2 s_3 \hat{b}_2' \mathrm{e}^{-s_3 \hat{b}_3' i}\, i}{\left(1 + s_2 \hat{b}_2' \mathrm{e}^{-s_3 \hat{b}_3' i}\right)^2} \\
\frac{\partial^2 r'(i)}{\partial \hat{b}_1' \partial \hat{b}_3'} &= \frac{2 s_1 s_2^2 s_3 \hat{b}_1' \hat{b}_2' e^{-2 s_3 \hat{b}_3' i} i}{\left(1 + s_2 \hat{b}_2' e^{-s_3 \hat{b}_3' i}\right)^3} - \frac{s_1 s_2 s_3 \hat{b}_1' e^{-i s_3 \hat{b}_3'} i}{\left(1 + s_2 \hat{b}_2' e^{-s_3 \hat{b}_3' i}\right)^2}
\end{aligned}
\tag{2.44}
$$

As done for the unscaled version of this example, substitute the partial derivatives of (2.43) and (2.44) into each term (2.37) of the Hessian matrix (2.33) to get an analytic expression that can be used to numerically calculate the Hessian. Using the actual parameters (Table 2.4) for the s_i values, the Hessian at the actual parameter location is calculated. The Hessian and its eigenstructure at the actual parameters are given in Table 2.6, where the scaling has clearly transformed the eigenstructure. All eigenvalues are now within two orders of magnitude of each other and the eigenvectors have changed direction.

Table 2.6 Hessian, eigenvalues, and eigenvectors for logistic equation example at actual parameter location with and without parameter scaling.

	Scaled			**Unscaled**		
	Hessian (N=1000)			**Hessian (N=1000)**		
	2.54e+05	-8.83e+04	2.30e+05	2.54e–01	–8.83e+00	5.76e+04
	-8.83e+04	3.80e+04	-8.42e+04	–8.83e+00	3.80e+02	–2.11e+06
	2.30e+05	-8.42e+04	2.17e+05	5.76e+04	–2.10e+06	1.36e+10
	Eigenvalue			**Eigenvalue**		
	λ_1	λ_2	λ_3	λ_1	λ_2	λ_3
	4.99e+05	7.19e+03	3.12e+03	1.36e+10	5.36e+01	9.62e–03
	Eigenvector			**Eigenvector**		
Axis	X_1	X_2	X_3	X_1	X_2	X_3
b_1	0.7099	0.5561	-0.4322	0.0000	–0.0020	–1.0000
b_2	-0.2558	0.7753	0.5775	–0.0002	–1.0000	0.0020
b_3	0.6562	-0.2994	0.6926	1.0000	–0.0002	0.0000

Parameters b_1 and b_3 have about the same contribution to the eigenvector with the largest eigenvalue, so we continue looking at these two parameters in 2-dimensional space. The contour curves of the cost function in the b_1–b_3 space are shown in Figure 2.6. This example shows that parameter scaling changes both the Hessian condition number and the importance of parameters as determined by their associated eigenvalues. Therefore, from here forward the parameters will always be scaled to a magnitude of 1 for computation of the Hessian. This includes all the analysis of the process-based Yield-SAFE model in Chapter 3.

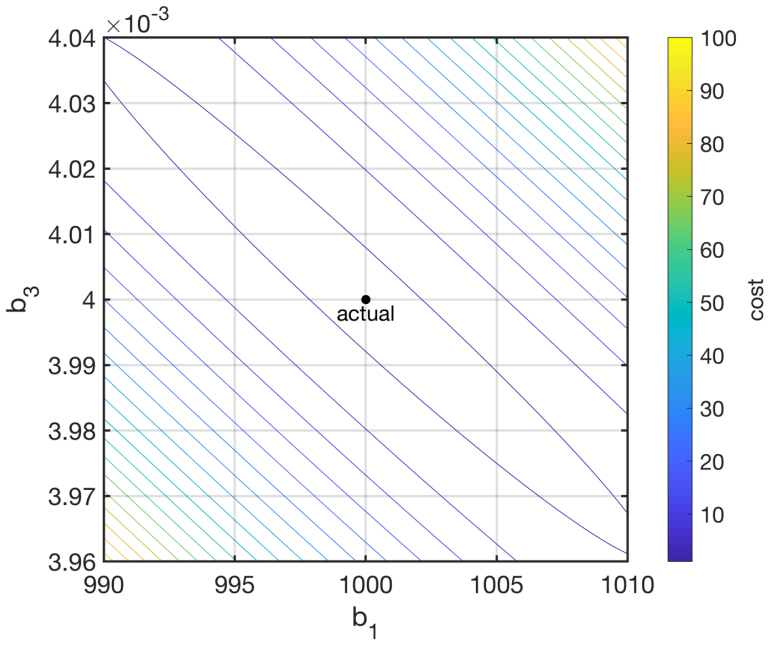

Figure 2.6 Contour lines for the logistic equation b_1–b_3 space with parameter scaling close to the actual parameters (±1% of parameter values).

The observation that all the Hessian terms are dependent upon the parameter values suggests that one should explore the Hessian some distance away from the actual parameters. In real-life problems, our *a priori* parameter estimates might be 5% or more away from the actual values. The contours of the cost function for ±20% of their actual b_1–b_3 values are shown in Figure 2.7. Here some interesting behavior becomes apparent, namely that some distance away from the actual parameters the contours become nonconvex. Even though the two-dimensional cost function has a single minimum (for the 251 x 251 grid of costs used to generate Figure 2.7), one side is nonconvex (concave downward) some distance from the origin. Therefore, the Hessian has both positive and negative eigenvalues in at least some regions of the parameter space.

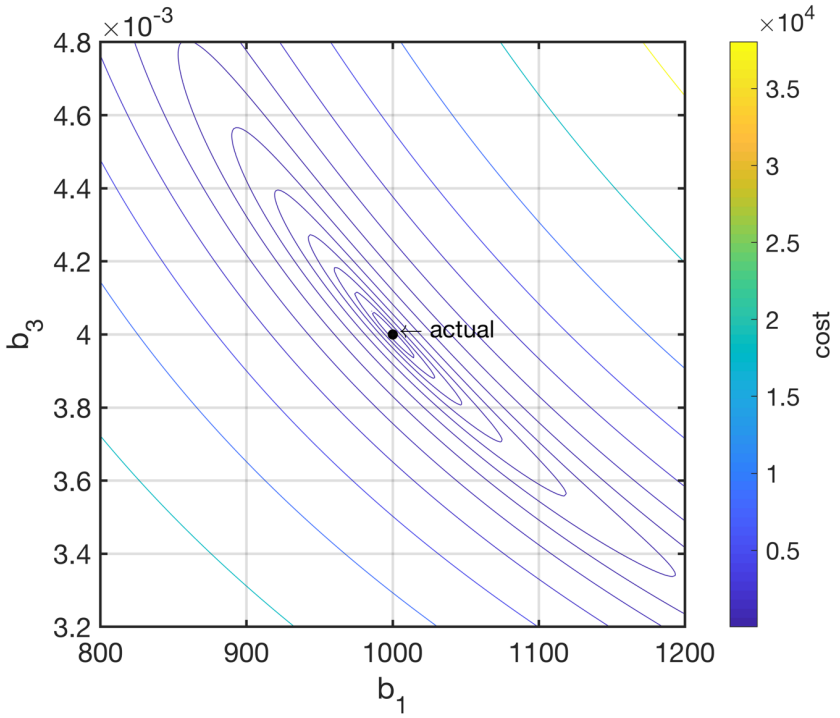

Figure 2.7 Contour lines for logistic equation in b_1–b_3 space at ±20% of the actual parameter values.

In most practical plant growth modeling one does not have the luxury of knowing the actual parameter values. For process-based models, there are a number of parameters that represent biophysical quantities that are known to greater or lesser degree. Therefore, the analysis of the logistic equation continues here from the point of view that one starts with a good guess of the parameter values, the *a priori* estimate. This example utilizes an *a priori* estimate where all parameters are –5% away from their actual values (Table 2.4). The output of the model with the *a priori* parameters as compared to the actual or reference output is shown in Figure 2.8. Because this investigation continues with Hessian analysis from *a priori* estimates, the term 'origin' will refer to the actual parameter location.

The goodness-of-fit measure given by the normalized root mean squared error (NRMSE) is used to compare the fit of each estimated output with the reference output (MathWorks 2018)

$$NRMSE\left(y(n)\right) \equiv \left(1 - \frac{\sqrt{\sum_{i=1}^{n}\left(y_0(i) - y(i)\right)^2}}{\sqrt{\sum_{i=1}^{n}\left(y_0(i) - \frac{\sum_{j=1}^{n} y_0(j)}{n}\right)^2}} \right) \times 100\% \tag{2.45}$$

where $y_0(i)$ is the reference output and $y(i)$ is the comparison output. The possible range of NRMSE values is 100% (perfect fit) to -∞. The threshold NRMSE value for an acceptable fit depends upon the context and, ultimately, user preferences. A user may consider a >90% fit is good enough in a certain setting, while another situation calls for a >99% fit.

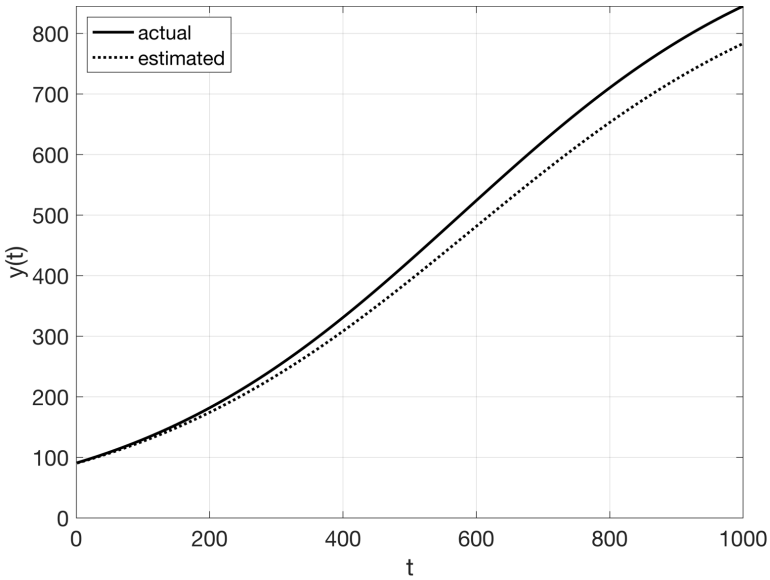

Figure 2.8 Output of logistic equation with actual and *a priori* estimated parameters. The NRMSE goodness-of-fit measure is 83.87%.

This *a priori* analysis scenario leads to a cost function that is more representative of what a modeler would encounter during the parameter estimation phase of model development. In general, the modeler has only good initial estimates; in this sense, it is much more realistic to investigate the Hessian away from the parameter origin. The Hessian for this *a priori* case is given in Table 2.7, where it

can be seen that eigenstructure has changed significantly from that of the origin, particularly in the direction of the eigenvectors. This is not surprising, considering the convex shape of the cost function at the actual parameters changes to nonconvex away from the origin. The condition number of the Hessian is 3e–3, which is reasonably small, suggesting that all parameters can successfully be estimated. However, we continue with the investigation of reducing the parameter space for estimation of the logistic equation parameters.

Table 2.7 Hessian, eigenvalues, and eigenvectors for logistic equation example at *a priori* parameter estimate and at origin.

	At *a priori* estimate			**At actual parameters**		
	Hessian (N=1000)			**Hessian (N=1000)**		
	2.17e+05	-7.33e+04	1.82e+05	2.54e+05	-8.83e+04	2.30e+05
	-7.33e+04	2.99e+04	-7.98e+04	-8.83e+04	3.80e+04	-8.42e+04
	1.82e+05	-7.98e+04	2.03e+05	2.30e+05	-8.42e+04	2.17e+05
	Eigenvalue			**Eigenvalue**		
	λ_1	λ_2	λ_3	λ_1	λ_2	λ_3
	4.22e+05	2.91e+04	-1.29e+03	4.99e+05	7.19e+03	3.12e+03
	Eigenvector			**Eigenvector**		
Axis	X_1	X_2	X_3	X_1	X_2	X_3
b_1	0.6913	-0.7216	0.0377	0.7099	0.5561	-0.4322
b_2	-0.2659	-0.2055	0.9419	-0.2558	0.7753	0.5775
b_3	0.6719	0.6611	0.3339	0.6562	-0.2994	0.6926

As seen in Figure 2.9, the contour lines of the cost function at the *a priori* parameters are similar to those of Figure 2.7, with the most obvious difference being that the minimum point has shifted away from the actual parameters. The shift is due to the position of b_2; in other words, if only b_1 and b_3 are estimated, their optimized values will differ from their true values as shown in the 3-parameter linear model example.

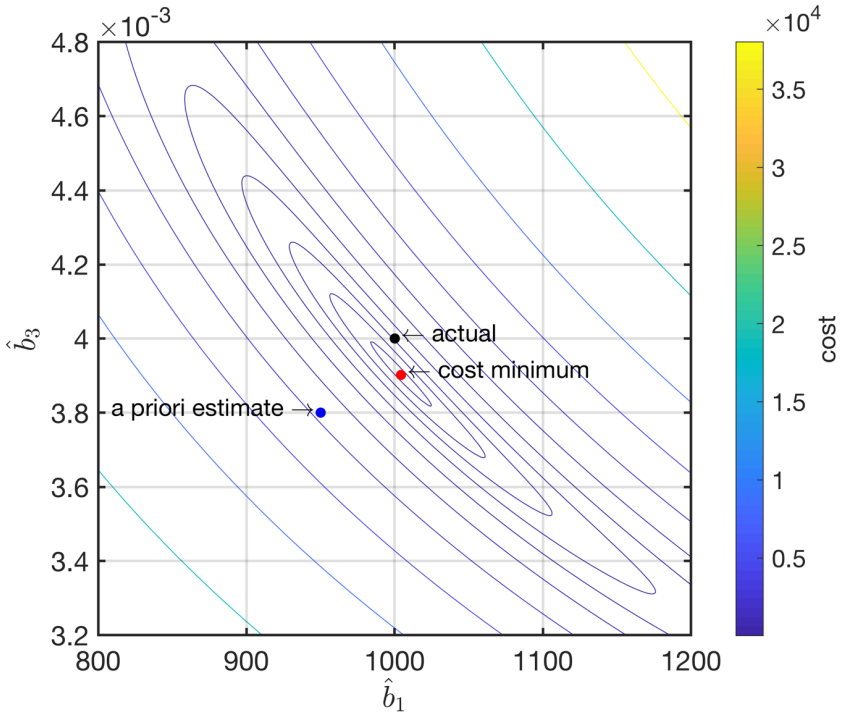

Figure 2.9 Contour lines for logistic equation in the $\hat{b}_1 - \hat{b}_3$ space with the *a priori* parameters.

In analysis of more complex nonlinear systems one would like to know that the eigenstructure at any point (our *a priori* estimate) in the feasible space is more-or-less representative of any point in that space in terms of which parameters are most important to the output dynamics. To get a sense of how the Hessian eigenstructure changes in $\hat{b}_1 - \hat{b}_3$ space, the Hessians along two transects through the contours depicted in Figure 2.7 are calculated along lines passing through each of the $\hat{b}_1$ and $\hat{b}_3$ *a priori* estimates (while holding the other two parameters at their *a priori* values). The three eigenvalues associated with each eigenvector are then plotted as a function of $\hat{b}_1$ or $\hat{b}_3$ at $\pm 20\%$ of their actual values. Recall that the magnitudes of the eigenvalues are of interest, indicating the curvature of cost function along their associated eigenvectors.

This analysis is shown in Figure 2.10 and Figure 2.11, confirming that although the magnitude of each eigenvalue changes over the ±20% range, their relative ranking by magnitude stays the same throughout the parameter space. This result lends some confidence to the notion that the Hessian can be used to rank parameters' importance to the curvature of the cost function in the nonlinear case.

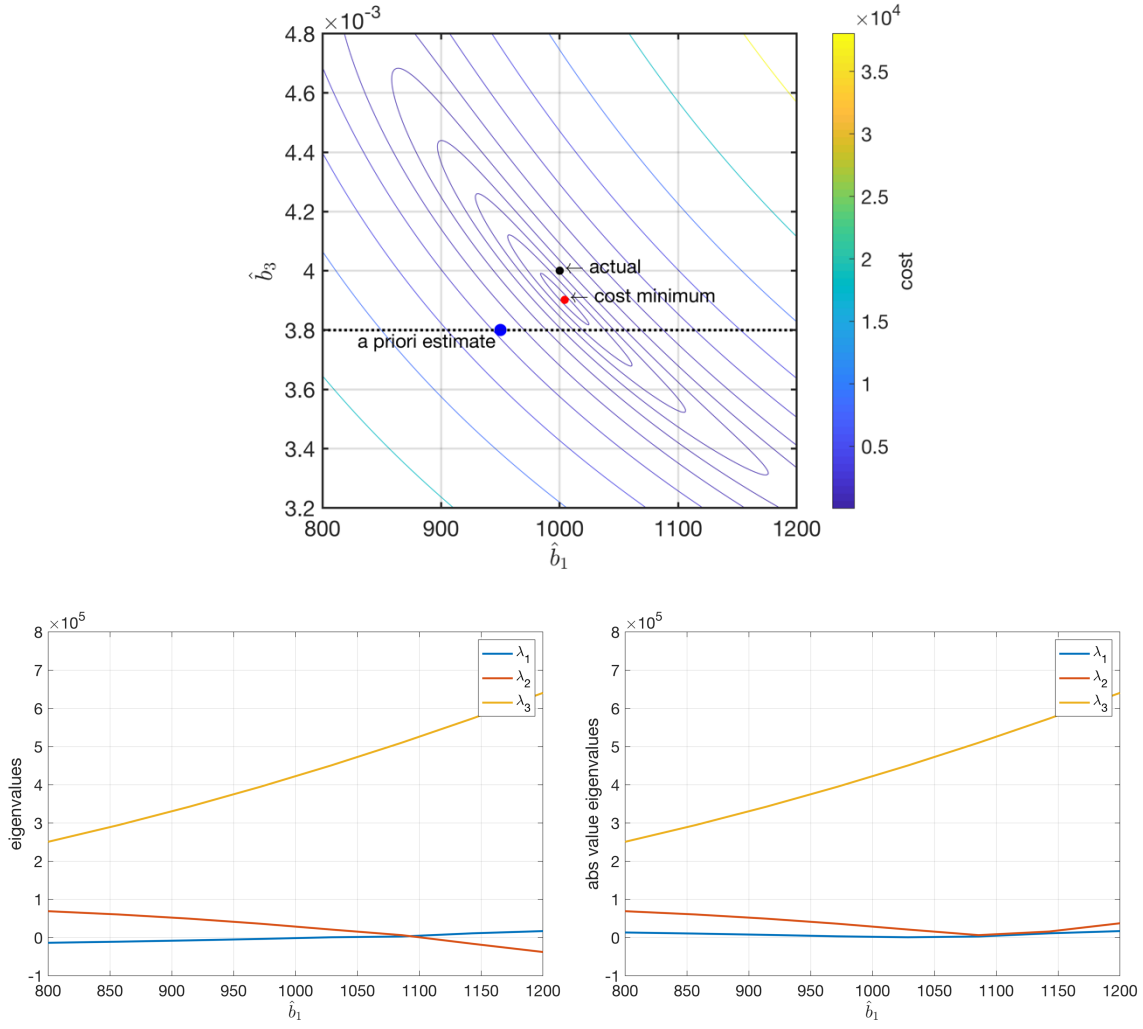

Figure 2.10 Change in eigenvalues associated with the eigenvectors X_i while varying $\hat{b}_1$ (along dotted line in top figure) while the other parameters are held at their *a priori* values. Figure on bottom left shows the eigenvalues, while the figure on bottom right shows the magnitude of the eigenvalues.

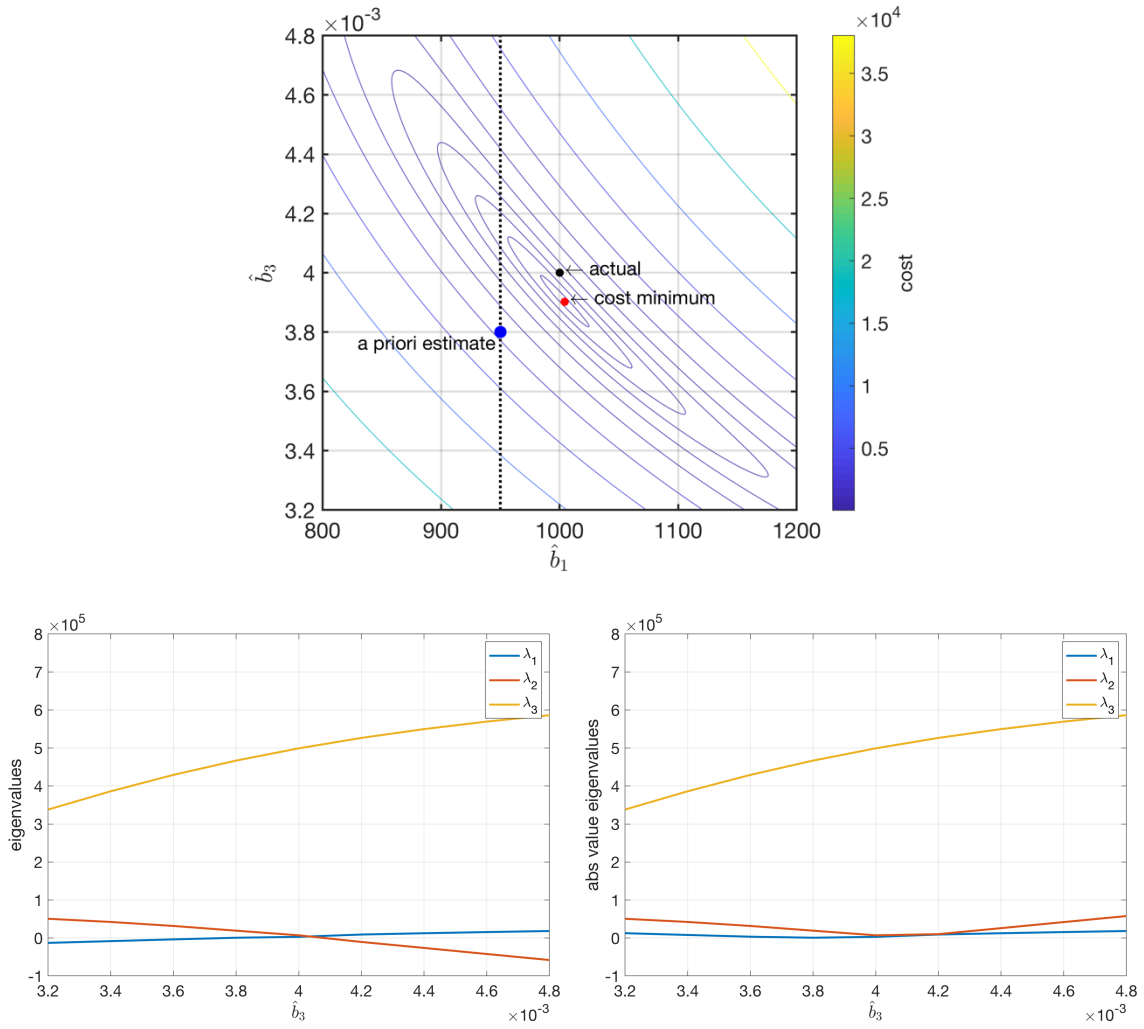

Figure 2.11 Change in eigenvalues associated with the eigenvectors X_i while varying $\hat{b}_3$ (along dotted line in top figure) while the other parameters are held at their *a priori* values. Figure bottom left shows the eigenvalues, while the bottom right figure shows the magnitude of the eigenvalues.

Now with some confidence in the Hessian analysis in a portion of the parameter space where one expects the optimum parameter setting is located, a method to order the importance of the parameters to the cost function curvature is needed. Those that have most influence on the curvature are the first candidate for optimization. As noted in the linear case, convergence along an eigenvector is dependent upon its eigenvalue (Widrow and Stearns 1986). However, the influence of individual parameters on the cost function curvature is of primary, in

order to select a subset of parameters to estimate. Therefore, more than the magnitude of the eigenvalues, we are interested in the projections of the eigenvectors onto the parameter axes multiplied by their associated eigenvalues. In this example, the projections of X_1 onto $\hat{b}_1$ and $\hat{b}_3$ are about the same length (0.6913 and 0.6719, respectively, from Table 2.7). In order to decide which of these is most important one might also consider the projections of the next eigenvector X_2 on to the parameter axes. Even though λ_1 is an order of magnitude bigger than λ_2, the projections of X_2 might help distinguish between the importance of $\hat{b}_1$ and $\hat{b}_3$. This reasoning could be applied to all eigenvectors in the case where the projections onto parameter axes and/or the eigenvalues are close in magnitude.

Based on this argument, the measure for ranking the importance of parameters to the Hessian curvature is assigned to be the 1-norm of the projections of the eigenvectors onto each parameter axis multiplied by their respective eigenvalues, that is,

$$\begin{aligned} \Lambda_i &= \left\| D\mathbf{X}_i^T \right\|_1 \\ &= \left\| \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} X_i^T \end{bmatrix} \right\|_1 \end{aligned} \tag{2.46}$$

where Λ_i is the combined magnitude of the eigenvector projections for the i^{th} parameter and X_i^T is the i^{th} column of the transpose of the eigenvector matrix. The Λ_i values at the *a priori* setting are shown in Table 2.8, which show that $\hat{b}_1$ and $\hat{b}_3$ are close in ranking. In an LMS estimation context, the Λ_i values correspond to the largest magnitude projection of eigenvectors possible along the

i^{th} parameter axis. As a parameter estimate approaches the a value that minimizes the cost function, the step along its axis will approach 0. However, when far away from the minimum, one could argue that Λ_i values are a reasonable indication of the relative cost function curvature along each parameter axis.

Table 2.8 Λ values at the *a priori* setting.

Parameter	Λ
$\hat{b}_1$	3.60e+05
$\hat{b}_2$	1.35e+05
$\hat{b}_3$	3.32e+05

Ideally, the relative importance of the parameters to the curvature of the cost function stays relatively constant. In other words, ideally the highest ranked parameters would be the same throughout the parameter space under consideration. To get a sense of this similar to the above calculations of the relative eigenvalues shown in Figure 2.11, Λ_i is calculated for locations along a line running through each of the $\hat{b}_1$ and $\hat{b}_3$ *a priori* estimates with the other parameters held at their *a priori* values. The results of this transect analysis are shown in Figure 2.12. Notice here that $\hat{b}_1$ and $\hat{b}_3$ have similar Λ values about 2–4 times greater than that of $\hat{b}_2$. This suggests that $\hat{b}_1$ and $\hat{b}_3$ have a bigger influence on the cost function than $\hat{b}_2$. Also, as an important side note, there is a point at which the relative importance of $\hat{b}_1$ and $\hat{b}_3$ switches (around $\hat{b}_1 = 950$ and $\hat{b}_3 = 0.0037$). This is a reminder that the analysis applies to the vicinity around the *a priori* parameter set and may have different results in other locations of the parameter space.

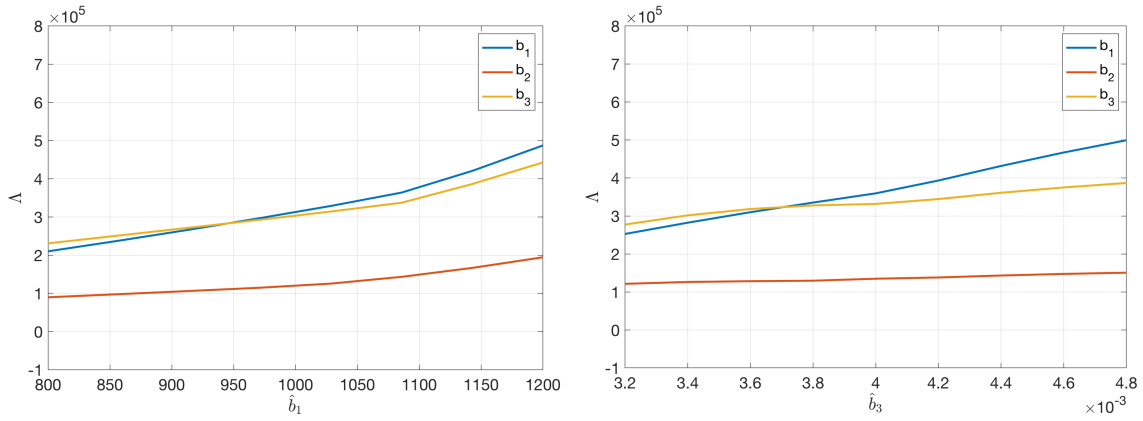

Figure 2.12 Λ values along lines passing through $\hat{b}_1$ (left) and $\hat{b}_3$ (right) while the other parameters are held at their *a priori* values.

Now with candidate parameters for estimation, $\hat{b}_1$ and $\hat{b}_3$, one can run a series of optimization trials for the candidate reduced-space parameter sets to determine goodness-of-fit (Table 2.9). Optimization was carried out using MATLAB's `lsqnonlin` function using the `trust-region-reflective` algorithm and then calculating a fit measure, the NRMSE. Optimizing on the parameter with highest ranking $\hat{b}_1$ initialized at its *a priori* value (and the others fixed at their *a priori* values) gives a much improved fit, as shown in Figure 2.13. Optimizing on $\hat{b}_3$ and both $\hat{b}_1$ and $\hat{b}_3$ lead to a similar goodness-of-fit (NRMSE >98%). While optimizing on the lowest ranking parameter $\hat{b}_2$ improves the fit (from 83.87% without parameter estimation to 91.90%), there is considerable room for improvement (Figure 2.14). As a final test, all three parameters are estimated and the converge to their actual values, suggesting that for this model, there is no reason to estimate a reduced number of parameters.

In conclusion, the information contained in the Hessian has led to a reasonable ranking of parameters in terms of their importance for optimization in the case of the three parameter logistic equation. At this location in the parameter space of

the logistic equation estimating $\hat{b}_1$ or $\hat{b}_3$ or both give a very good model fit. In the next chapter, this analysis is applied to a much more complex nonlinear model.

Table 2.9 Optimization of various parameters and combinations of parameters initiated from the *a priori* values.

Trial	$\hat{b}_1$	$\hat{b}_2$	$\hat{b}_3$	NRMSE
Actual parameters	1000	10	0.004	100%
A priori values (no optimization)	950	9.5	0.0036	83.87%
Optimize $\hat{b}_1$ only	1028.8	9.5	0.0036	98.14%
Optimize $\hat{b}_2$ only	950	7.9753	0.0036	91.90%
Optimize $\hat{b}_3$ only	950	9.5	0.004145	97.06%
Optimize $\hat{b}_1$ & $\hat{b}_3$	1003.7	9.5	0.003947	98.44%
Optimize $\hat{b}_1$ & $\hat{b}_2$	1038.8	9.7522	0.0036	98.38%
Optimize $\hat{b}_2$ & $\hat{b}_3$	950	10.146	0.004250	97.76%
Optimize $\hat{b}_1$, $\hat{b}_2$ & $\hat{b}_3$	1000	10	0.004	100%

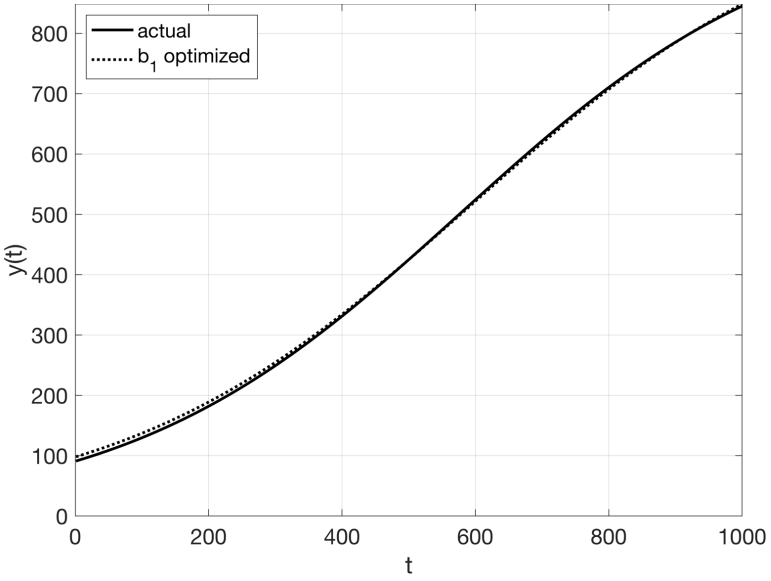

Figure 2.13 Output of *a priori* logistic model with $\hat{b}_1$ optimized. The NRMSE goodness-of-fit measure is 98.14%.

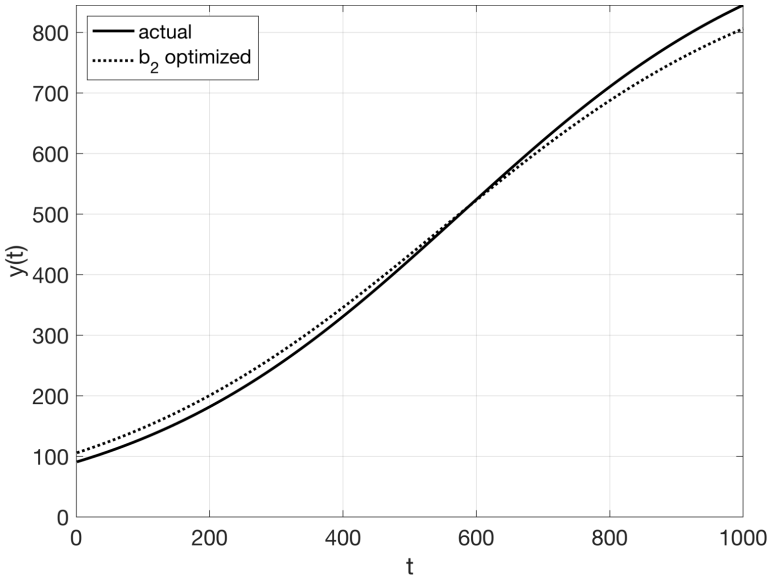

Figure 2.14 Output of *a priori* logistic model with $\hat{b}_2$ optimized. The NRMSE goodness-of-fit measure is 91.90%.

2.7 The Hessian-informed Reduced-Order Parameter Estimation (HIROPE) procedure

The previous linear and nonlinear examples lead to a procedure for determining a reduced-order space for parameter estimation for deterministic nonlinear growth models.

2.7.1 Model prerequisites

The procedure begins with a process-based model that is considered a promising representation of the real system, the full system model. The goal is to reduce the model to a lower dimensional parameter space by fixing parameter values and estimating others that are consistently identifiable.

If successful, the resulting reduced-order parameter space for estimation, based upon the data from one field trial, will yield a parameterization that can be trusted to give good output prediction for any input within the input class. The procedure described here may not be the best method of determining the

reduced-order parameter space, however, the example case shown in Chapter 4 demonstrates that the approach is feasible. The primary objective of this dissertation is to persuade skeptics within the tree growth modelling community of the feasibility of an approach of this nature.

Application of the procedure has the following prerequisites

a. Full deterministic process equations are defined.
b. All inputs and outputs are readily measurable.
c. Determine *a priori* conditions on parameters, including mathematical and 'realism' constraints. May be set by modelers (e.g., all parameters are positive real), or by model itself (e.g., to follow realism constraints such as prohibiting occurrence of negative biomass).
d. The model is implemented on a computational platform.
e. If model equations are not twice continuously differentiable with respect to the parameters of interest, acceptable differentiable analytic approximations are available.
f. General class of model inputs is known and ample input data at the model time step can be accessed.
g. Acceptable *a priori* parameter values and initial conditions are available.
h. Cost function is defined.
i. Measure of fit function is defined, and a cut-off criterion for goodness-of-fit is selected by the user based on model goals.
j. A measure of consistent identifiability for parameters is defined.
k. A parameter optimization algorithm is implemented.

A list of symbols used in this analysis is given in Table 2.10.

Table 2.10 Symbols used for models, parameterizations, parameter initializations, and inputs with notes related to Yield-SAFE simulations in next chapter.

Symbol	Description	Notes
M	The model as described by the modelers	Models for tree (M_t), crop(M_c), tree/crop combination (M_{tc}), denote Yield-SAFE models
N		Number of time steps in simulation (days for Yield-SAFE)
Θ	Vector of model parameters considered for estimation (symbols)	Dimension 1 x d. $d = 6$ Tree & crop models (water non-limiting) or $d =$ 12 (combination tree/crop water nonlimiting)
θ_0	Actual or true parameter values	Dimension 1 x d. $d = 6$ Tree & crop models (water non-limiting) or $d =$ 12 (combination tree/crop water nonlimiting)
θ_j	*A priori* parameter values with each element located at +5% or –5% of its actual value	Dimension 1 x d. $j = 1, 2, \ldots, 2^d$
$\hat{\theta}_j$	*A priori* parameter values with some or all parameters optimized and others fixed at their θ_j values	Dimension 1 x d. $j = 1, 2, \ldots, 2^d$
$\hat{\hat{\theta}}_j$	*A priori* parameter values that meet model fit criterion with some or all parameters optimized and others fixed at their θ_j values	Dimension 1 x d. $j = 1, 2, \ldots, 2^d$
L_i	Locations	Locations selected for study. Index i is location number. 10 locations used in simulations.
u_i	Inputs	First column is solar radiation (MJ m^{-2}) and second is temperature (°C). Index i is associated with location L_i. A 2 x N array.
y_0	Actual or reference model output(s)	Biomass (g $tree^{-1}$ for M_t; g m^{-2} for M_c; both outputs for M_{tc}), with θ_0 and reference input u_i, a 1 x N sequence for M_t and M_c, and 2 x N for M_{tc}.

y_j	*A priori* output(s)	Generated with $\boldsymbol{\theta}_j$ and reference input u_i, a 1 x N sequence for M_t and M_c, and 2 x N for M_{tc}.
$\hat{y}_j$	*A priori* output(s) with some or all parameters optimized	Generated with $\hat{\boldsymbol{\theta}}_j$ and reference input u_i, a 1 x N sequence for M_t and M_c, and 2 x N for M_{tc}.
$\hat{\hat{y}}_j$	*A priori* output(s) that meet model fit criterion with some or all parameters optimized	Generated with $\hat{\hat{\boldsymbol{\theta}}}_j$ and reference input u_i, a 1 x N sequence for M_t and M_c, and 2 x N for M_{tc}.
V	Loss function $V_N(u_i, \boldsymbol{\theta}_j, y_0, y_j)$	Value computed from analytical expression using u_i, $\boldsymbol{\theta}_j$, y_0, and y_j over N steps. A scalar quantity.
$H_{i,j}$	Hessian of loss function $H(V_N(u_i, \boldsymbol{\theta}_j, y_0, y_j))$	Derived from analytical expression of $V_N(u_i, \boldsymbol{\theta}_j, y_0, y_j)$. Values computed over N steps. Dimension is d x d.
X_k	Eigenvectors of $H_{i,j}$	Dimension is d x 1.
λ_k	Eigenvectors of $H_{i,j}$	A scalar.
Λ_k	Ranking value for parameter k	Computed from $H_{i,j}$ by ranking parameters by the 1-norm of the eigenvector projections scaled by their respective eigenvalues onto each parameter axis. A scalar quantity.
$\chi_{j,k}$	Candidate reduced parameter sets (k) for optimization from *a priori* parameter set $\boldsymbol{\theta}_j$	Computed from parameter rankings for all parameters in Θ. Dimension is d x d.

2.7.2 Steps of the HIROPE procedure

The steps of the proposed model reduction procedure are outlined here.

Step 1. Compare goodness-of-fit of y_0 with the estimated output y_j.

For one u_i, compare output y_0 for a parameter set $\boldsymbol{\theta}_0$ with the output y_j for

one of the associated *a priori* θ_j parameter set using the NRMSE criterion (Figure 2.15). If the fit meets the criterion, move to Step 6 (validation).

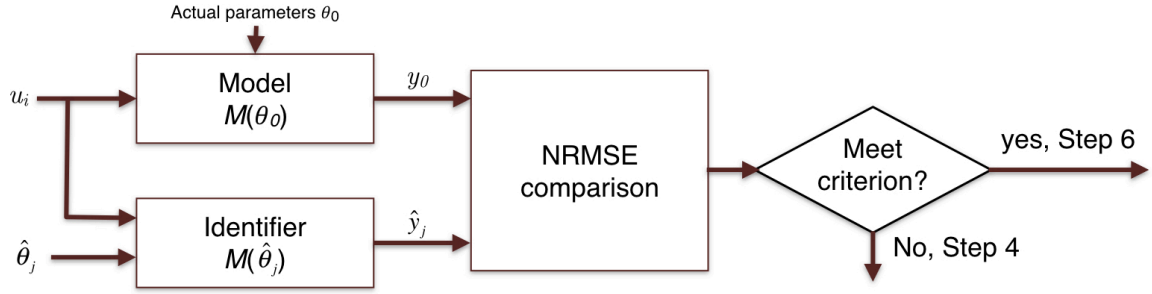

Figure 2.15 Procedure Step 1.

Step 2. Compute $H_{i,j}$ and its eigenvectors and eigenvalues. Compute $H_{i,j}$ for θ_j using data sequences u_i, y_0, and y_j from above simulation and determine its eigenvectors and eigenvalues (Figure 2.16).

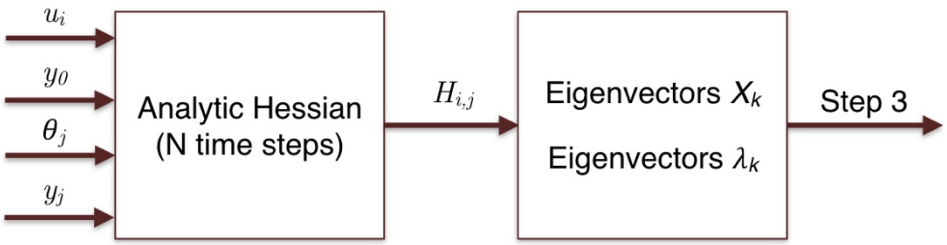

Figure 2.16 Procedure Step 2.

Step 3. Generate Λ parameter ranking. From the computed $H_{i,j}$, rank parameters in order of highest to lowest magnitude of associated eigenvalues using the formula (2.46) (Figure 2.17). In cases where the projection of an eigenvector is equal in magnitude for two or more parameters, choose one of them, then both in the ranking.

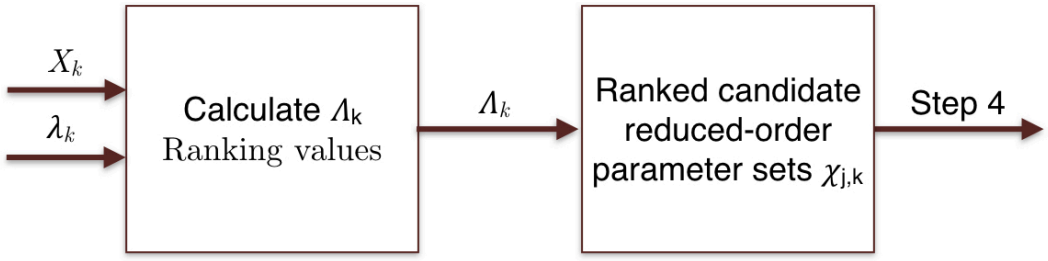

Figure 2.17 Procedure Step 3.

Step 4. Confirm consistent identifiability. Beginning with highest ranked reduced-order parameter set $\chi_{j,1}$ containing only one parameter associated with the largest eigenvalue, optimize the parameter value from both of its ±5% values (with all other parameters in θ_j fixed to their original estimated values) to determine a $\hat{\theta}_j$ (Figure 2.18).

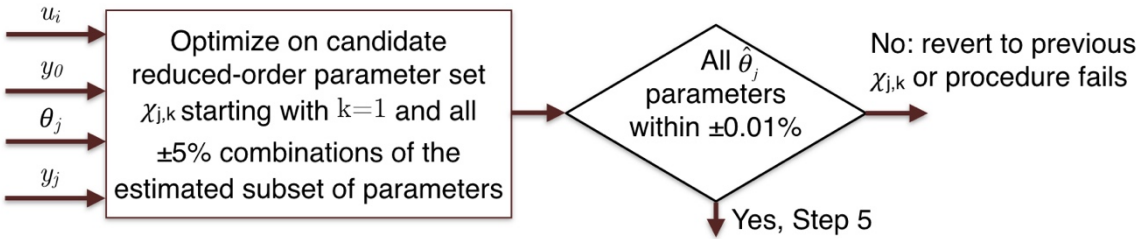

Figure 2.18 Procedure Step 4.

Step 5. Check goodness-of-fit of y_0 and $\hat{y}_j$. Using the current $\hat{\theta}_j$, generate $\hat{y}_j$ and calculate the NRMSE value relative to y_0 (Figure 2.19). If criterion (NRMSE ≥99%) is met, then move to the next step for validation. If optimization on the first candidate parameter set does not meet the NRMSE criterion, then repeat Step 4 for the next candidate parameter set, and so on, until either the fitting criterion is met for the optimization and validation inputs (success) or not (failure of procedure). In case of failure, revise fit criterion or model.

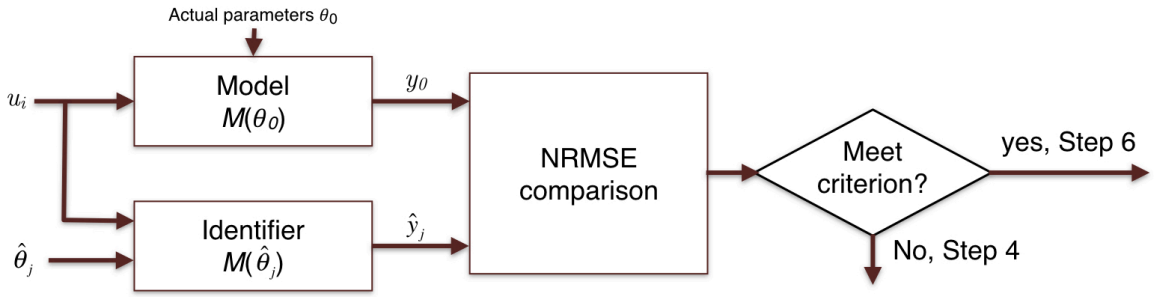

Figure 2.19 Procedure Step 5.

Step 6. Validation. Validate the optimized parameter set using a number of different input sequences from the same location and calculating their NRMSE values (Figure 2.20). If the validating input data also results in an acceptable NRMSE value, then the model optimization procedure is considered successful, otherwise return to Step 4. One may proceed to the next candidate parameter set if one wishes to see if there is improvement.

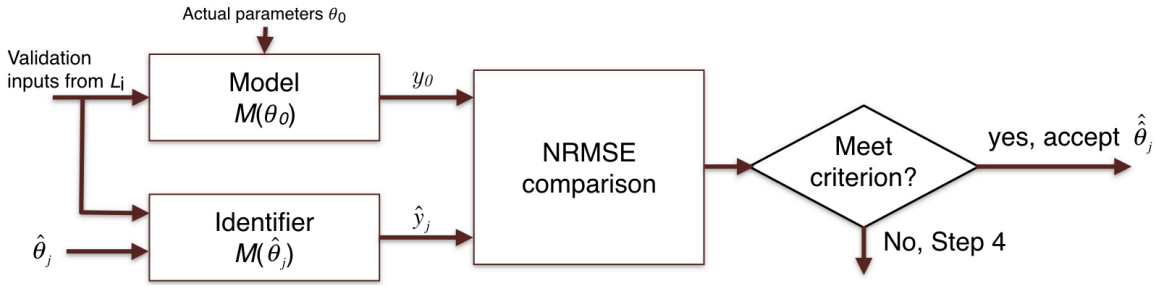

Figure 2.20 Procedure Step 6.

The full HIROPE procedure is illustrated in Figure 2.21.

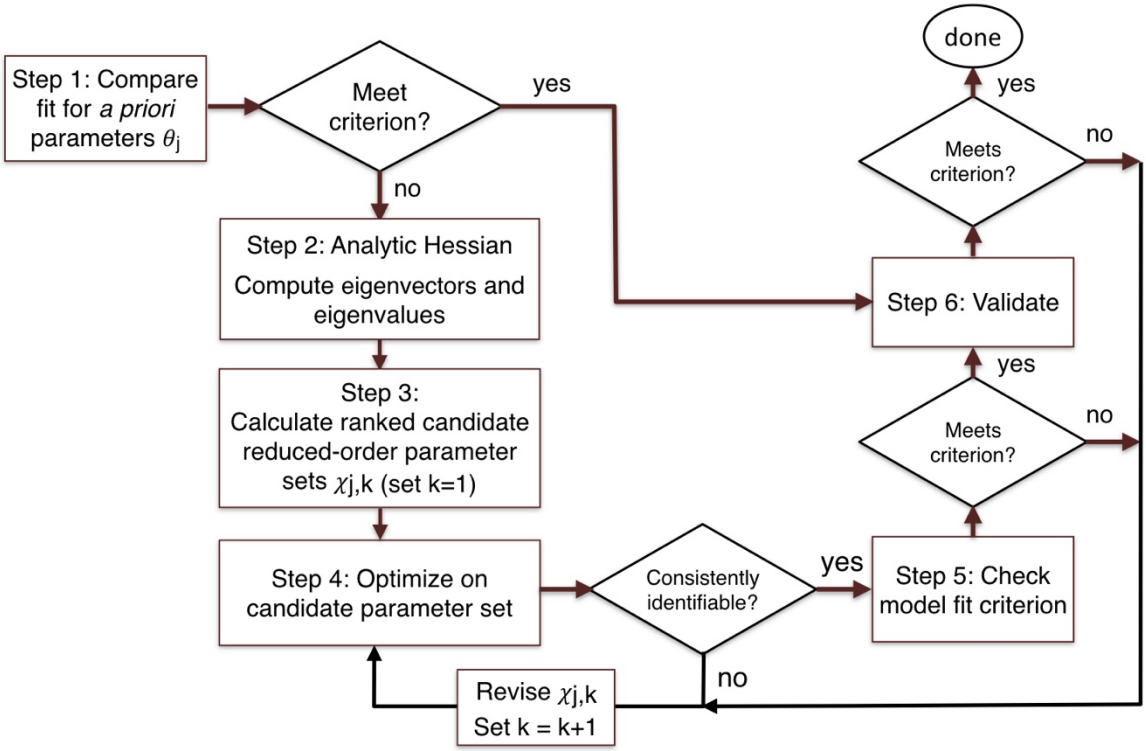

Figure 2.21: The 6-step Hessian-informed reduced-order parameter estimation (HIROPE) procedure.

3 APPLICATION OF THE HIROPE PROCEDURE TO YIELD-SAFE MODEL

This chapter demonstrates step-by-step the successful application of the parameter space reduction procedure presented in Chapter 2 applied to the Yield-SAFE model (van der Werf et al. 2007).

Covered in this chapter are:

1. Yield-SAFE model background
2. Preliminary analysis of Yield-SAFE
    a. Model fit from *a priori* parameter sets
    b. Parameter identifiability test
    c. One-at-a-time parameter sensitivity
3. Step-by-step application of the procedure from Section 2.7 to Yield-SAFE tree, crop, and combination tree-crop models with accompanying analysis

The next chapter presents an expanded range of procedure simulations and accompanying meta-analysis for the Yield-SAFE model.

3.1 The Yield-SAFE model

The Yield-SAFE model was developed with lower complexity than many other tree growth models. The seminal paper on Yield-SAFE emphasizes its reduced complexity (van der Werf et al. 2007) in its title, "Yield-SAFE: A parameter-sparse, process-based dynamic model for predicting resource capture, growth and production in agroforestry systems." Developed by the Silvoarable Agroforestry for Europe (SAFE) project during the 2000's, the Yield-SAFE model estimates biomass yields of tree rows integrated with arable crops (Dupraz et al. 2005, van

der Werf et al. 2007, Graves et al. 2010). The model consists of process-based descriptions of tree and crop growth in a two-story planting configuration driven by environmental inputs (solar radiation, temperature, and precipitation), i.e., it is an input-output model that includes both plant growth and interactions between trees and crops (Figure 3.1 and Figure 3.2). Recently, Yield-SAFE was augmented with several new submodels, including vapor pressure deficit input (to predict transpiration), modified water uptake, and the effect of the trees on temperature and wind speed, among others (Palma et al. 2016). As an initial case study, the analysis below is based upon the original version of Yield-SAFE described in van der Werf et al. (2007) and implemented in Microsoft Excel (Burgess et al. 2014). The analysis is expected to be applicable to the recently introduced more complex versions of Yield-SAFE as well as other process-based models that meet the procedure criteria.

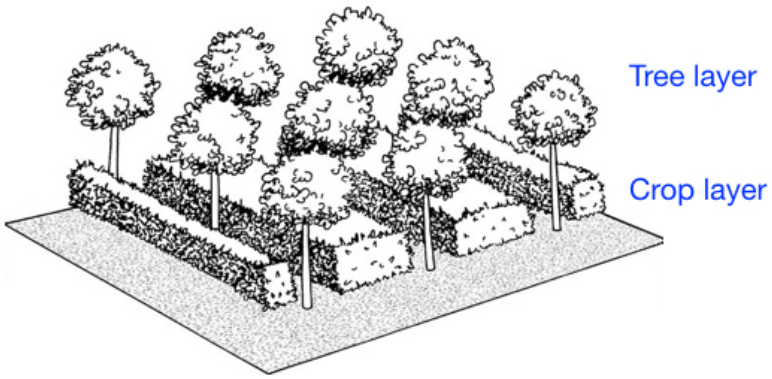

Figure 3.1: Tree and understory crop configuration for Yield-SAFE predictive model. (Illustration by Christi A. Sobel.)

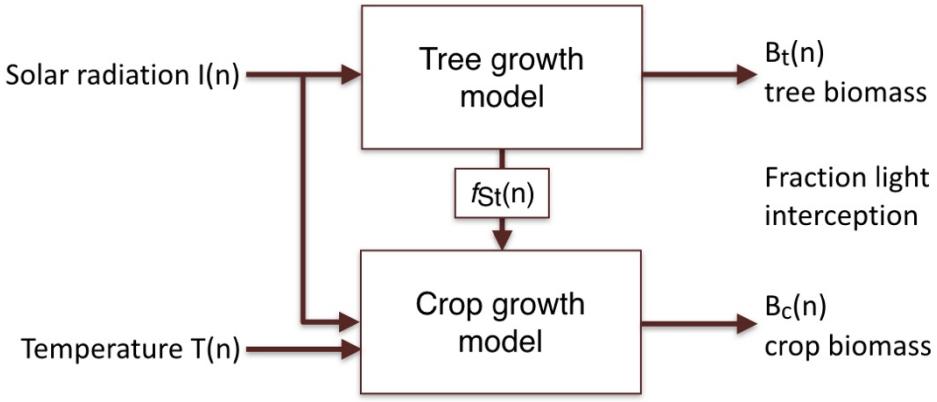

Figure 3.2: Yield-SAFE block diagram in water non-limiting configuration.

Yield-SAFE is unique in that it (1) is the only widely known tree growth model with understory crops developed specifically for yield estimation, (2) has been evaluated for a range of crops over the past decade or so, and (3) has released example parameterizations and initial conditions for various trees and crops (Burgess et al. 2014, Palma et al. 2017). The model's intended predictive ability serves several objectives including yield forecasts, management scenario testing, economic projections and optimization, and background for agricultural policy decisions.

Additionally, as a process-based or mechanistic model, Yield-SAFE is unusual in that it was designed to be less complex with fewer parameters than many other process-based tree growth models. As stated in van der Werf et al. (2007), "The model was developed with as few equations and parameters as possible to allow model parameterization under constrained availability of data from long term experiments." In general, given sufficient numbers of input-output data, a simpler model leads to less uncertainty in parameter estimates and model predictions (Ljung 1999). Even though Yield-SAFE was developed with a minimal modeling philosophy, an initial investigation of the model's performance when parameters

are only approximately known demonstrates that parameter estimation is necessary (Section 3.1.4).

Reducing the number of signals that are of interest within a system identification context opens possibilities for further model reduction. Rather than model the behavior of internal states of tree growth that cannot easily be measured (e.g., leaf area, number of branch shoots), in many applications attention can be limited to readily measurable outputs of interest (e.g., a measure of biomass such as trunk diameter or height). The number of parameters estimated can be reduced if variations in some parameters (or combinations of parameters) contribute little to the model output(s). The degree of contribution is a combination of spectral complexity of the input and parameter sensitivities of the model structure (Yao et al. 2003, Gutenkunst et al. 2007, Crout et al. 2009, Li and Vu 2013).

3.1.1 Model description

The Yield-SAFE model under consideration here is described in detail in van der Werf et al. (2007). The model includes seven state equations (1) tree biomass; (2) tree leaf area; (3) number of shoots per tree; (4) crop biomass; (5) crop leaf area index; (6) soil water content; and (7) heat sum. The model runs on a daily time step and is driven by exogenous inputs of solar radiation, temperature, and precipitation. For the purposes of the present analysis, water and soil nutrients will be considered sufficient and non-limiting to potential growth. This simplification to streamline the initial demonstration is based upon Liebig's Law of the Minimum, a simplification traditionally applied to crop growth, which states that plant growth is limited by the scarcest resource (van Ittersum and Rabbinge 1997, van der Ploeg 1999). The water and nutrient non-limiting

assumption is consistent with parameter estimation done by the Yield-SAFE team (van der Werf et al. 2007, Keesman et al. 2011, Palma et al. 2017). The trees are also assumed to be healthy, i.e., negative growth (dieback or death) due to long-term drought, severe nutrient deficiency, or pest and disease are not considered here.

Although this assumption is made for the purposes of model analysis, it is not realistic in terms of real-life crop growth. Relaxation of these simplifying assumptions still needs to be tested on the Yield-SAFE model expanded with additional submodules such as soil water and nutrient dynamics, both of which increase complexity significantly by adding numerous connections between the crop and tree submodels. However, it is anticipated that the concepts applied here are scalable to models with larger number of parameters and more complex dynamics. It is assumed that the model reduction concepts demonstrated here under the water and nutrient non-limiting assumption will next be applied in future research to the full model including water and nutrient submodels.

According to Dupraz et al. (2005), the Yield-SAFE model was originally implemented in a MATLAB/Simulink environment, but was later implemented in Excel. The growth equations used here are based upon the Excel implementation by Burgess et al. (2014) without management interventions such as tree thinning and pruning (or the water and soil modules, as previously mentioned). The Excel version is based upon the Yield-SAFE equations given in van der Werf et al. (2007), Graves et al. (2010), and Keesman et al. (2011).

3.1.1.1 Tree growth model

The Yield-SAFE tree growth model in the non-water-limiting case has a single input (solar radiation) and no interactions with the understory crop that

influence tree growth. Therefore, the tree growth model can be run independently of the crop growth model.

The modeled rate of tree growth is given by the product of the proportion of incoming radiation intercepted by the trees, the radiation use efficiency, and the quantity of incoming solar radiation ($I(n)$ (MJ m^{-2})). The proportion of radiation intercepted is a function of the number of branch shoots and the leaf area. Initial conditions are biomass and number of shoots and for the crop leaf area. A full derivation of the tree growth model from a mechanistic perspective is given in van der Werf et al. (2007). In this analysis, the system equations from the Excel implementation (Burgess et al. 2014) are used, which have some revisions as compared with the van der Werf et al. (2007) description, such as the addition of a "phased light extinction coefficient" described in Keesman et al. (2011).

$B_t(n)$ is tree biomass (gm/tree) at time n (days from planting)

$$B_t(n) = B_t(n-1) + \frac{\varepsilon_t}{\rho} I(n)\left(1 - e^{-\rho k(n) L_t(n)}\right) - \alpha B_t(n-1) \tag{3.1}$$

$N(n)$ is the number of shoots per tree

$$N(n) = N(n-1) + (L_m R)^{1/2} \left(\frac{\frac{B_t(0)}{N(0)}\left((L_m R)^{1/2} - 1\right)}{\left(B_t(n) + \frac{B_t(0)}{N(0)}\left((L_m R)^{1/2} - 1\right)\right)^2} \right) \left(B_t(n) - B_t(n-1) \right) \tag{3.2}$$

$L_t(n)$ is leaf area (m^2/tree)

$$L_t(n) = \left(L_t(n-1) + \left(1 - e^{-\frac{1}{\tau}}\right)\left(\left(\frac{L_m}{R} \right)^{1/2} N(n-1) - L_t(n-1) \right) \right) \phi(n) \tag{3.3}$$

where $\phi(n)$ (dimensionless) is given by

$$\begin{aligned} &\phi(n) = 1 \quad \text{DOYbudb} \le \text{DOY} \le \text{DOYleaff} \\ &\text{else} \\ &\phi(n) = 0 \end{aligned} \tag{3.4}$$

and $k(n)$ is light extinction (dimensionless)

$$k(n) = k_t - k_b + k_b\left(\frac{k_a}{L_t(n) + k_a}\right) \tag{3.5}$$

The tree model (3.1)–(3.5) is depicted graphically in Figure 3.3. $L_t(n)$ is set to zero in (3.4) during a winter period defined by day-of-year bud burst (DOYbudb) and day-of-year leaf fall (DOYleaff). Day-of-year (DOY) is incremented daily beginning at 0 on January 1 and is reset to 0 the following January 1. The phenological parameters DOYbudb and DOYleaff are considered to be fixed, known values for the purposes of the analysis. Equation (3.4) sets leaf area and therefore the growth portion (middle term) of (3.1) to zero for a period of time, after which leaves regrow and biomass growth commences again.

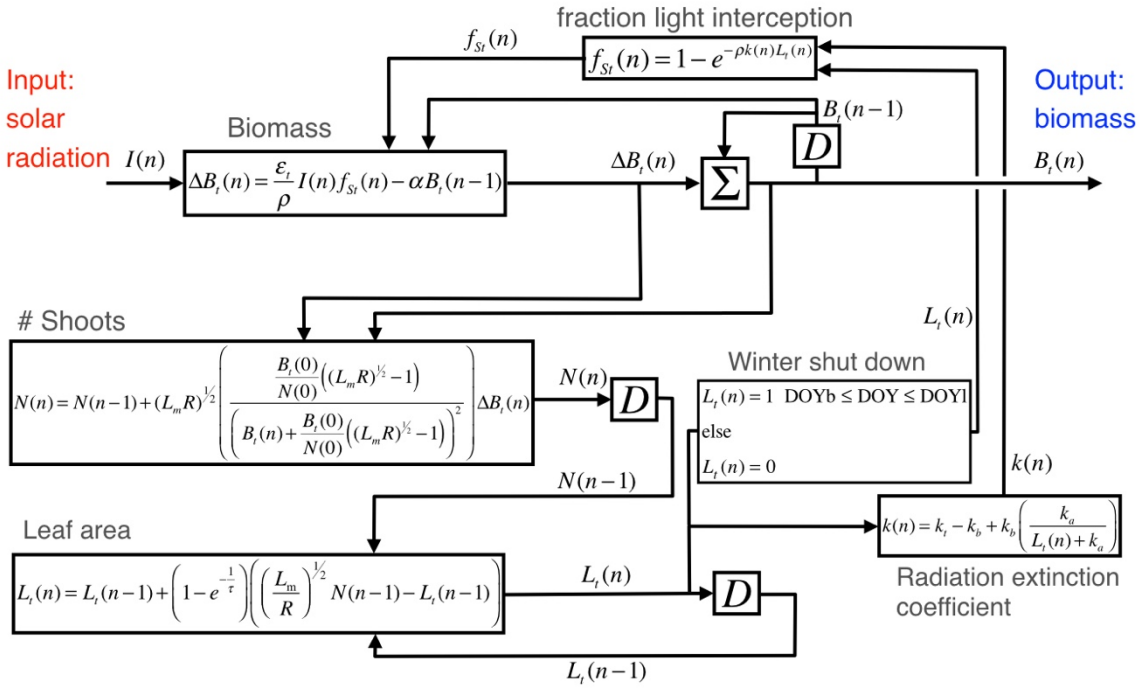

Figure 3.3: Yield-SAFE tree growth (overstory) model diagram based upon (3.1)–(3.5). This diagram depicts the non-water-limiting case used for analysis.

3.1.1.1.1 Tree parameters

Model parameters and initial conditions are positive real (Table 3.1). The symbol ρ is tree density per m^2, which is based upon planting density, considered a known user-assigned constant for the simulation. Table 3.2 lists the tree parameterizations included with the Excel implementation (Burgess et al. 2014). For the purposes of the model analysis conducted here, these six parameterizations are assumed to be deemed by the Yield-SAFE project as 'accepted,' i.e., they lead to model outputs that closely match tree growth, and therefore are assigned to be the actual parameters θ_0 for model analysis.

Table 3.1 Parameters used in Yield-SAFE tree growth model in the non-water limiting case (Burgess et al. 2014).

Symbol	Description	Units
ε_t	Radiation use efficiency	g MJ^{-1}
α	Attrition rate of standing biomass	day^{-1}
L_m	Maximum leaf area	m^2
$B_t(0)$	Initial biomass per tree	gm $tree^{-1}$
N(0)	Initial number of shoots per tree	$tree^{-1}$
R	'Ratio' related to leaf and shoot maxima	–
τ	Time constant of leaf area growth	days
k_t	Light extinction coefficient	–
k_a	Light extinction "a" coefficient	–
k_b	Light extinction "b" coefficient	–
DOYbudb	Day-of-year of bud burst	–
DOYleaff	Day-of-year of leaf fall	–
ρ	Tree stand density (user-assigned constant based on planting configuration)	trees m^{-2}

Table 3.2 Accepted actual parameter values given for various trees in Yield-SAFE Excel implementation (Burgess et al. 2014).

Symbol	Units	Poplar1 (Graves 2010)	Miscanthus	Wild cherry	AGF Poplar2011	FOR Poplar2011	Apple (5.7 x 2.8 m)
ε_t	g MJ^{-1}	1.4086	0.9	0.5626	1.05	0.9	0.5626
α	day^{-1}	0.0001	0.0004	0.0001	0.0003	0.0003	0.00005
L_m	m^2	500	15	500	240	200	150
$B_t(0)$	gm $tree^{-1}$	100	55	55	55	50	80
N(0)	$tree^{-1}$	0.6225	0.6225	0.5713	0.6225	0.7	3
R	–	200000	100	200000	50000	240000	200000
τ	days	10	10	10	10	10	10
k_t	–	0.8	0.8	0.8	0.8	0.8	0.8
k_a	–	10	10	10	10	10	10
k_b	–	0.4	0.4	0.4	0.4	0.4	0.4
DOYbudb	–	100	100	100	100	100	135
DOYleaff	–	300	300	300	300	300	310
ρ Forestry	trees m^{-2}	0.1089	0.1089	0.1089	0.1089	0.1089	0.1089
ρ Agroforestry	trees m^{-2}	0.0156	0.0156	0.0156	0.0156	0.0156	0.0156

3.1.1.2 Crop growth model

As for tree growth, crop growth is modeled based on light interception by the crop and radiation use efficiency. The understory growth has two inputs, (1) solar radiation $I(n)$ modified by the amount of light intercepted by the overstory $f_{St}(n)$ and (2) temperature $T(n)$. The crop growth descriptive equations are based upon van der Werf et al. (2007) as implemented in Burgess et al. (2014)

$$B_c(n) = B_c(n-1) + (1 - f_{St}(n))I(n)\varepsilon_c\left(1 - \mathrm{e}^{-k_c L_c(n-1)}\right) \tag{3.6}$$

where $B_c(n)$ is crop biomass ($\mathrm{gm/m^2}$)

$$L_c(n) = L_c(n-1) + \Delta B_c(n)P(n)\sigma \tag{3.7}$$

where $L_c(n)$ is crop leaf area ($\mathrm{m^2\ m^{-2}}$)

$$\Delta B_c(n) = B_c(n) - B_c(n-1) \tag{3.8}$$

$P(n)$ is partitioning to leaf (dimensionless)

$$\begin{aligned} P(n) &= P_0 && S(n) \le S_1 \\ P(n) &= P_0 \frac{S_2 - S(n)}{S_2 - S_1} && S_1 < S(n) \le S_2 \\ P(n) &= 0 && S(n) > S_2 \end{aligned} \tag{3.9}$$

and $S(n)$ is temperature sum (°C-days)

$$S(n) = S(n-1) + \max[0, T(n) - T_0] \tag{3.10}$$

The positive real model parameters are ε_c, T_0, P_0, S_1, S_2, σ, and k_c (with $S_2 > S_1$, for units and biophysical interpretations see Table 3.3). The block diagram in Figure 3.4 represents equations (3.6)–(3.10).

In the Excel implementation (Burgess et al. 2014), additional parameters are included, as listed in Table 3.3 and Table 3.4. Parameters DOYsow and DOYharv are management-selected days of the year for sowing and harvesting, respectively. S_0 is the temperature sum to crop emergence, while S_h is the temperature sum for harvest (unless DOYharv has been exceeded). Equations (3.6)–(3.10) apply only during the crop growth period between DOYsow and DOYharv, with growth restarting from near zero on an annual periodic basis.

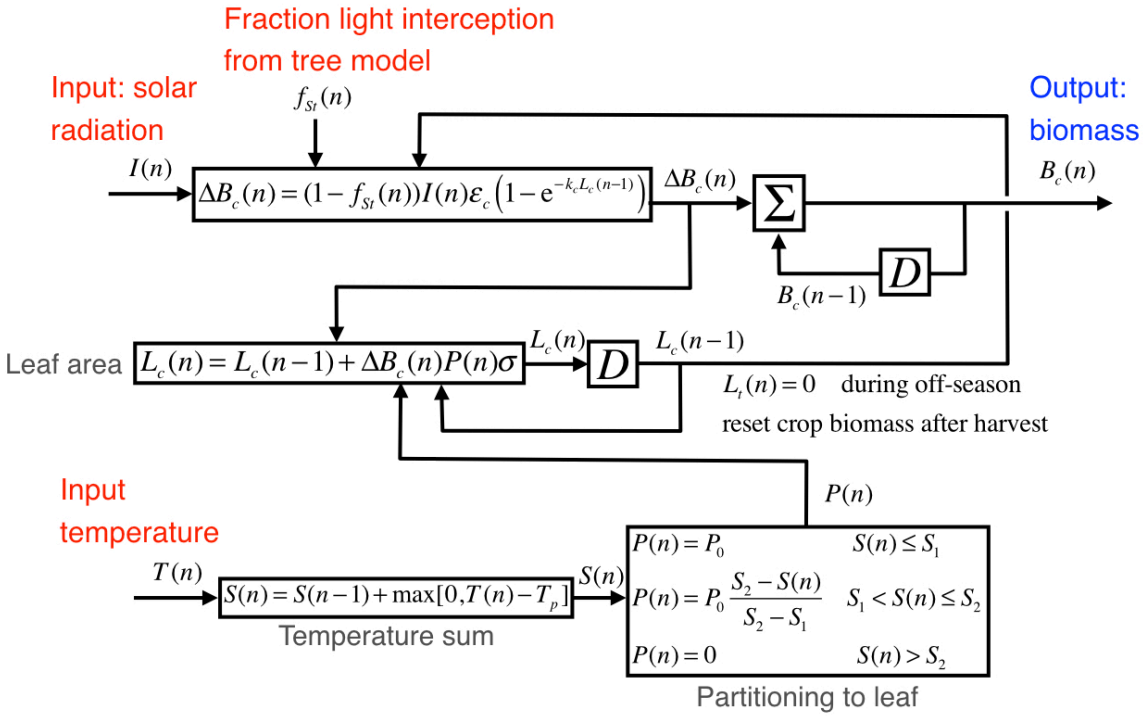

Figure 3.4: Yield-SAFE crop (understory) growth model based on van der Werf et al. (2007).

Table 3.3 Crop model parameters and initial conditions for the crop in the non-water limiting case (Burgess et al. 2014).

Symbol	Description	Units
ε_c	Potential growth	g MJ^{-1}
T_0	Temperature threshold	°C
P_0	Partition to leaves at emergence	–
S_1	T-sum at which partitioning starts to decline	°C days
S_2	T-sum at which partitioning to leaves = 0	°C days
σ	Specific leaf area	m^2 gm^{-2}
k_c	Radiation extinction coefficient	–
DOYsow	Day-of-year of sowing	–
DOYharv	Day-of-year of harvest (if threshold not reached)	–
S_0	Temperature sum to emergence	°C days
S_h	Temperature sum to harvest	°C days
$B_c(0)$	Initial crop biomass	g m^{-2}
$L_c(0)$	Initial leaf area	m^2 m^{-2}

Table 3.4 Crop model parameter values accepted for various understory crops in Yield-SAFE Excel implementation (Burgess et al. 2014).

Symbol	Units	WWheat	WBarley	SBarley	W Beans	S Field bean	Sugar beet	Annual grass	Forage maize
ε_c	g MJ^{-1}	1	1	1	1	1	1.2	0.3	1.4
T_0	°C	0	0	0	0	0	3	0	8
P_0	–	0.8	0.8	0.8	0.6	0.6	0.9	0.8	0.8
S_1	°C days	1600	1300	1300	1490	790	50	1600	250
S_2	°C days	1840	1500	1500	1500	800	680	1840	600
σ	m^2 gm^{-2}	0.021	0.021	0.021	0.02	0.02	0.017	0.021	0.021
k_c	–	0.4	0.45	0.45	0.5	0.5	0.73	0.4	0.4
DOYsow	–	-107	-107	60	-75	60	93	-107	115
DOYharv	–	235	230	235	257	257	312	243	283
S_0	°C days	150	150	150	100	100	50	150	150
S_h	°C days	2800	2440	1770	2500	1900	2080	3200	960
$B_c(0)$	g m^{-2}	0	0	0	0	10	0.3	0	10
$L_c(0)$	m^2 m^{-2}	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

3.1.1.3 Tree-crop combination model

The combination model in the water non-limiting case is described by the equations for the tree and crop model presented above. All the parameters of the

tree and crop models carry over to the combination model, with no additional parameters. Interaction between the two models is the fraction of light intercepted by the trees, which is subtracted from the daily quantity of incoming solar radiation impacting the crop. See Figure 3.5 for depiction of the combination model.

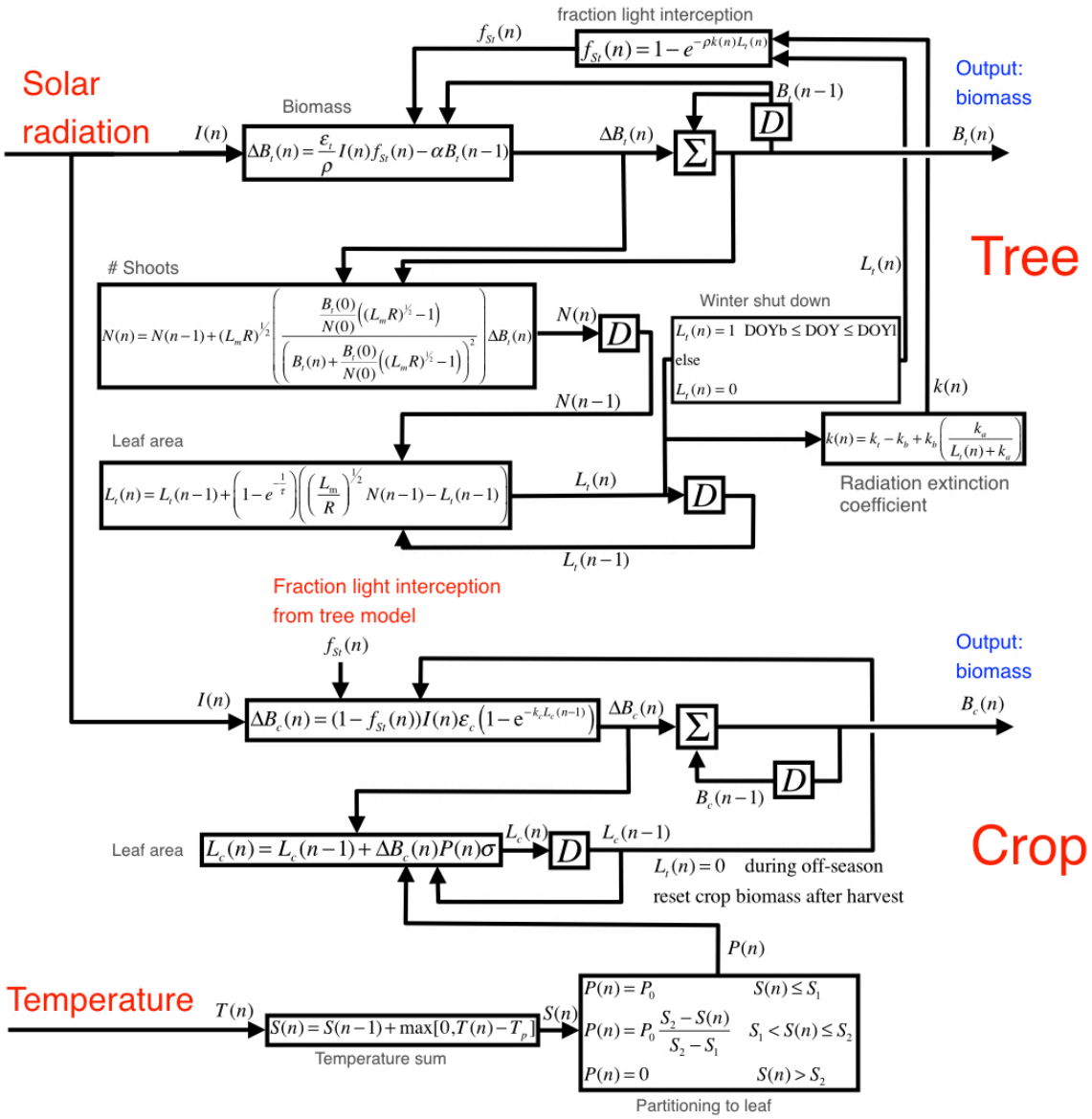

Figure 3.5 Tree-crop combination model.

3.1.2 Yield-SAFE project's parameterization methods

Due to lack of I/O data for two-story agroforestry configurations, the Yield-SAFE developers have parameterized the model based on yield tables, other models (such as STICS, see Brisson et al. 2003), and field data for monocultures (single-species plantings) of the tree and crop species of interest. Van der Werf et al. (2007) described how parameters are determined through a two-step process

1) Curve fitting to accepted yield tables for tree monocultures and accepted model yields for arable crop monocultures (using least squares optimization and manual adjustments).
2) Fine-tuning of certain parameters for a specific site based upon monoculture growth data.

According to van der Werf et al. (2007), two parameters are adjusted for the tree model: the initial number of branch shoots $N(0)$ and radiation use efficiency ε_t. For crops, radiation use efficiency ε_c, heat sum parameters (S_0, S_1, S_2, and S_h), and harvest index (HI_c) are adjusted. Notably, rather than parameterizing the combination model as one model, the Yield-SAFE parameterization steps include only tree and crop monoculture curve fitting to parameterize the full tree-crop combination model.

Keesman et al. (2011) approaches Yield-SAFE parameterization following "a constrained parameter optimization approach, where the constraints have been obtained from experts, expressed in terms of individual parameter ranges." Additional steps include, "... a good deal of expert judgement, taking into account both the plausibility of parameter values as described in the literature, and the

match between model results and yield levels at the landscape test sites." No specific, reproducible method for parameter estimation is given.

Palma et al. (2017) expands upon these earlier Yield-SAFE parameterization descriptions, as summarized here

- Gather experimental data, if available.
- Determine site latitude and longitude for acquiring daily real or simulated climate data as well as soil depth and texture.
- Note thinning and pruning management (timing and residual biomass).
- Check which system quantities are included in the Yield-SAFE model (e.g., leaf area per tree, fraction light interception, biomass per tree) to determine which observed data can be used for calibration.
- Align measured data to simulation calendar (count days since January 1st).
- Run the model against observed data.
- Plot the observed data against model predicted quantities.
- Calibrate/adjust parameters according to the physiological range of parameters from literature as much as possible. This can be done manually or by using parameter estimation software.

The final step assumes literature review into the plausible physiological range of parameters related to the biophysical submodel equations. Palma et al. (2017) further explain that initial calibration of the growth model is carried out "using only irrigated data and 'switching-off' the water module of the model", meaning that water is assumed to be non-limiting. After estimating the growth parameters, other data from control plots are used to calibrate the water submodel parameters. A suggested method for the non-water limiting portion of

the estimation procedure is optimization using the L-BFGS-B quasi-Newton algorithm (Byrd et al. 1995) on model parameters with their values constrained to upper and lower bounds found in the literature. The L-BFGS-B optimization module accompanies the Yield-SAFE model version implemented in Python (Palma et al. 2017). In the initial non-water limiting phase of the parameter estimation, observed data for biomass and leaf area are used. Palma et al. (2017) states, "The parameters for which there were values in the literature are set and the other ones are allowed to vary between biologically relevant values," in other words, only the parameters that are considered to be unknown are estimated. In the next step of the estimation, measured data for volume, height and diameter values are used in a manual calibration "changing each parameter value at a time." Finally, once the growth parameters are calibrated, control measurements are used to calibrate the water module parameters, "while fine-tuning the other parameters as a whole."

Palma et al. (2017, pp. 74–162) presents examples of this parameter estimation process for many trees and crops. These examples do not explicitly state which parameters are estimated, nor which parameters are fixed to values found in the literature, estimated through an optimization algorithm, or manually adjusted. A reproducible procedure is presented here for determining a lower dimension subspace for parameter estimation that may benefit model users who would like to parameterize the model for new crops and environments.

3.1.3 Inputs and Validation

Daily climatic data for solar radiation, temperature, and precipitation are required to run the original Yield-SAFE model (Table 3.5). Additional inputs including relative humidity and wind speed are required to drive the new Yield-

SAFE modules (Palma et al. 2016) for crop transpiration and tree effects on microclimate (temperature and wind), which are not considered here. Such data are often not available for a particular location or required time period of the simulation. Additionally, the Yield-SAFE model has been developed to project yields for possible future climate scenarios.

Table 3.5. Daily inputs to the original Yield-SAFE model (van der Werf et al. 2007, Burgess et al. 2014).

Symbol	Description	Unit/Value
$T(n)$	Temperature	°C on day n
$I(n)$	Solar radiation	MJ m^{-2} on day n
$R(n)$	Precipitation	m^3 m^{-2} on day n
Time		
Day	Day of the month	1–31
$Month$	Month of the year	Jan–Dec
$Year$	Year	xxxx
$Day\ of\ year$	Day of the year	Number 0–365, January 1 = 0, 365 only leap years

In order to acquire the needed data, the Yield-SAFE developers draw on simulated climate data to drive their model. Their current project (AGFORWARD) developed a web portal CliPick ("Climate Change Web Picker," http://home.isa.utl.pt/~joaopalma/projects/agforward/clipick/) to access climate data generated by a number of climate modeling projects (Palma 2015, Palma 2017). CliPick users enter the site longitude and latitude, the date range, time step (day or month), and select a data set. The data sets available through CliPick include a number of future climate scenarios as well as climate data simulated based upon historical measurements of atmospheric aerosols (van Meijgaard et al. 2012).

Palma et al. (2017) states, "There are indications that the simulated climate can be used for calibration purposes with minor loss of quality in comparison to real data." Palma et al. (2018) confirms, "The use of these datasets can certainly widen the usage of forest growth process based models, improving the support for decision-making in forest management."

To check the accuracy of the simulated inputs for particular location, CliPick recommends that it be compared with 20–30 year averages of real data. See Section 4.2 for more detail on input sequences accessed through CliPick that are used in this analysis.

Once the model parameters have been optimized on an input sequence and meet the model fit criteria (called $\hat{\hat{\theta}}_j$), the model parameterized by $\hat{\hat{\theta}}_j$ must be verified on other inputs with the same properties as the original input. Various data sequences extracted from the historical climate scenario were used for both the optimization and validation simulations conducted in this chapter (see Chapter 4 for details on sources of climate data used here). The basis for using historical data to simulate climate data is documented in Lamarque et al. (2011). Data sequences used for optimization and validation in this chapter are given in Table 3.6.

Table 3.6. Input sequence sources for simulations in this chapter. All data were imported through the CliPick portal.

Symbol	Location	Coordinates	Source	Start date	Length (days)
$u_{1,1}$	L_1, Norwich, UK	Lat: 52.6628 Long: 1.2283	Hist KNMI-RACMO22E	31-12-51	7306
$u_{1,2}$	Same	Same	Hist KNMI-RACMO22E	31-12-66	7306
$u_{1,3}$	Same	Same	Hist KNMI-RACMO22E	31-12-86	7306
$u_{1,4}$	Same	Same	A1B - HadCM3Q0	31-12-51	7306
$u_{1,5}$	Same	Same	A1B - HadCM3Q0	31-12-71	7306

3.1.4 Preliminary analysis of Yield-SAFE model

Before applying the model reduction procedure to Yield-SAFE, it is illuminating to conduct some exploratory simulations into model behavior. First, the model output error for a range of *a priori* parameter settings is considered. Second, an attempt is made to estimate all model parameters given a large amount of input-output data. Third, one-at-a-time parameter sensitivity in the output is considered as a means to determine a reduced-order parameter set for estimation.

3.1.4.1 Model output error from a priori *parameter values*

3.1.4.1.1 Tree growth model

One might ask how poor model performance would be if one has close approximations to the actual model parameters, but not their exact values. By 'close' a small percentage error is chosen, for this analysis $\pm 5\%$. For evaluation of the tree growth model, a set of actual parameters that are accepted is selected, as given in Table 3.1 ("Poplar1 (Graves 2010)"). These choices lead to a MATLAB implementation of the model with actual parameters and a given input ($u_{1,1}$) to calculate the or reference output y_0 (see Figure 3.6).

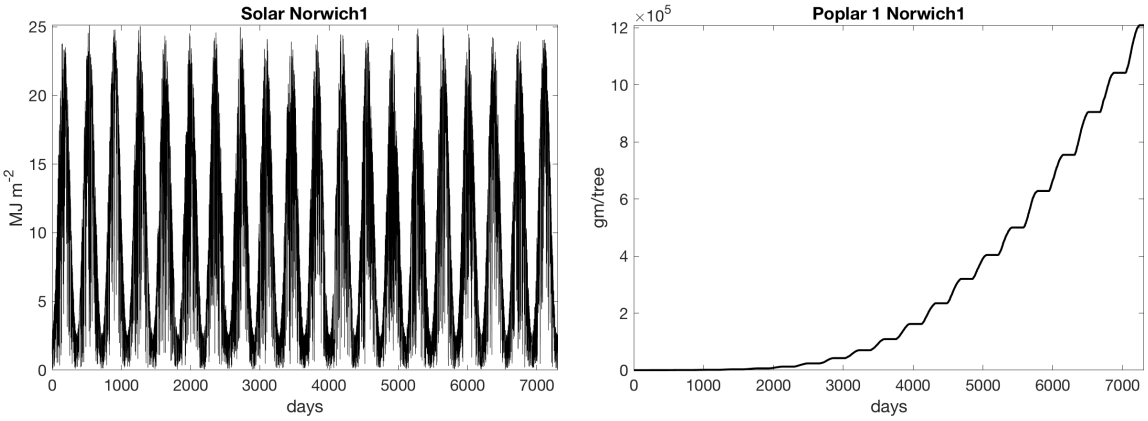

Figure 3.6. Solar radiation input $u_{1,1}$ (left), and tree growth model output y_0 (right) with parameters for Poplar1.

Next, we build a set of *a priori* parameter estimates in which all parameters are set to values either 5% above or below their actual values. Since values for τ, k_t, k_a, and k_b are the same for all crops in Table 3.1, these are assumed to be known, fixed constants, which also reduces the dimension of this example down to six parameters. For six parameters, there are 2^6 possible $\pm 5\%$ combinations, or 64 *a priori* parameter sets (Table 3.7). Each of these *a priori* sets generates a different realization of the output for a given input, as shown in Figure 3.7.

As seen in Figure 3.7, a histogram of the 64 NRMSE values shows approximately 70% of the *a priori* parameter settings lead to NRMSE values of $<90\%$ and 50% of parameter estimates give NRMSE values $<80\%$, model fits that most would consider unacceptable. This example demonstrates that some parameter adjustment must be done based on real data, even if the user is very skilled at estimating initial parameter values.

Table 3.7 The 64 θ_j *a priori* parameter settings for a 6-dimensional parameter vector. The symbol "–" indicates –5% away from actual parameter setting and "+" indicates +5% away from actual parameter setting.

	Parameter #							Parameter #					
θ_j	1	2	3	4	5	6	θ_j	1	2	3	4	5	6
tree	ε_t	α	L_m	$B_t(0)$	N(0)	R	tree	ε_t	α	L_m	$B_t(0)$	N(0)	R
crop	ε_c	T_0	P_0	S_1	S_2	k_c	crop	ε_c	T_0	P_0	S_1	S_2	k_c
1	–	–	–	–	–	–	33	–	–	–	–	–	+
2	+	–	–	–	–	–	34	+	–	–	–	–	+
3	–	+	–	–	–	–	35	–	+	–	–	–	+
4	+	+	–	–	–	–	36	+	+	–	–	–	+
5	–	–	+	–	–	–	37	–	–	+	–	–	+
6	+	–	+	–	–	–	38	+	–	+	–	–	+
7	–	+	+	–	–	–	39	–	+	+	–	–	+
8	+	+	+	–	–	–	40	+	+	+	–	–	+
9	–	–	–	+	–	–	41	–	–	–	+	–	+
10	+	–	–	+	–	–	42	+	–	–	+	–	+
11	–	+	–	+	–	–	43	–	+	–	+	–	+
12	+	+	–	+	–	–	44	+	+	–	+	–	+
13	–	–	+	+	–	–	45	–	–	+	+	–	+
14	+	–	+	+	–	–	46	+	–	+	+	–	+
15	–	+	+	+	–	–	47	–	+	+	+	–	+
16	+	+	+	+	–	–	48	+	+	+	+	–	+
17	–	–	–	–	+	–	49	–	–	–	–	+	+
18	+	–	–	–	+	–	50	+	–	–	–	+	+
19	–	+	–	–	+	–	51	–	+	–	–	+	+
20	+	+	–	–	+	–	52	+	+	–	–	+	+
21	–	–	+	–	+	–	53	–	–	+	–	+	+
22	+	–	+	–	+	–	54	+	–	+	–	+	+
23	–	+	+	–	+	–	55	–	+	+	–	+	+
24	+	+	+	–	+	–	56	+	+	+	–	+	+
25	–	–	–	+	+	–	57	–	–	–	+	+	+
26	+	–	–	+	+	–	58	+	–	–	+	+	+
27	–	+	–	+	+	–	59	–	+	–	+	+	+
28	+	+	–	+	+	–	60	+	+	–	+	+	+
29	–	–	+	+	+	–	61	–	–	+	+	+	+
30	+	–	+	+	+	–	62	+	–	+	+	+	+
31	–	+	+	+	+	–	63	–	+	+	+	+	+
32	+	+	+	+	+	–	64	+	+	+	+	+	+

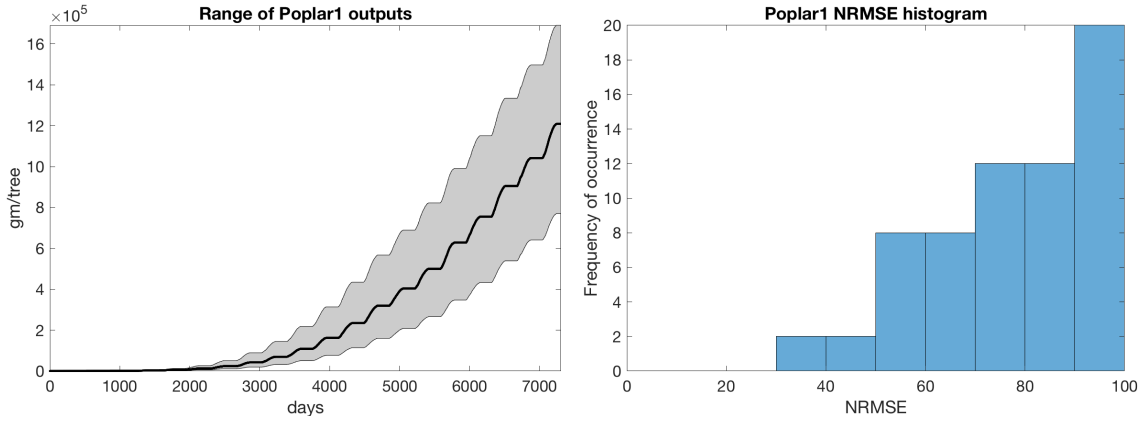

Figure 3.7. The range of tree growth (Poplar1) outputs for the 64 *a priori* parameter sets shown in the grey region, with the reference output in black (left). Histogram of NRMSE's for the 64 outputs (right). Model is driven by $u_{1,1}$.

3.1.4.1.2 Crop growth model

The same analysis for the understory crop model (in isolation from the tree crop model, that is, with unobstructed sunlight or $f_{St}(n) = 0$ in (3.6)) leads to similar results as for the tree growth model. To reduce the dimension of the understory model to six parameters, we first recognize that $P(n)$ multiples σ in (3.7). Since $P(n)$ is the product of P_0 and 0, 1, or $\frac{S_2 - S(n)}{S_2 - S_1}$, then (3.7) always contains the product of P_0 and σ. Because there is no way to uniquely identify the two parameters in a product, one of them (σ) is set to its actual value for this exercise and throughout this paper. Additionally, DOYsow, DOYharv, S_0, S_h and initial conditions $B_c(0)$ and $L_c(0)$ are assumed to be known and fixed values. This assumption is not considered to reduce the generality of the analysis, but renders the discussion more manageable for the purposes of demonstrating the challenges of parameter estimation and the potential usefulness of further reducing the number of parameters to be estimated.

The reference inputs and output together with the 64 are shown in Figure 3.8 for the crop parameters for "Annual grass" from Table 3.3. Figure 3.9 shows results similar to those of the tree growth model, with over 60% of the *a priori* parameter settings giving rise to NRMSE values <90%, and 25% leading to NRMSE values <80%, indicating that the crop model also requires parameter optimization even if the user begins with 'close' *a priori* parameters. In the combined model, where light interception by the tree model $f_{St}(n)$ influences crop growth rates, model performance will be worse that shown in Figure 3.9.

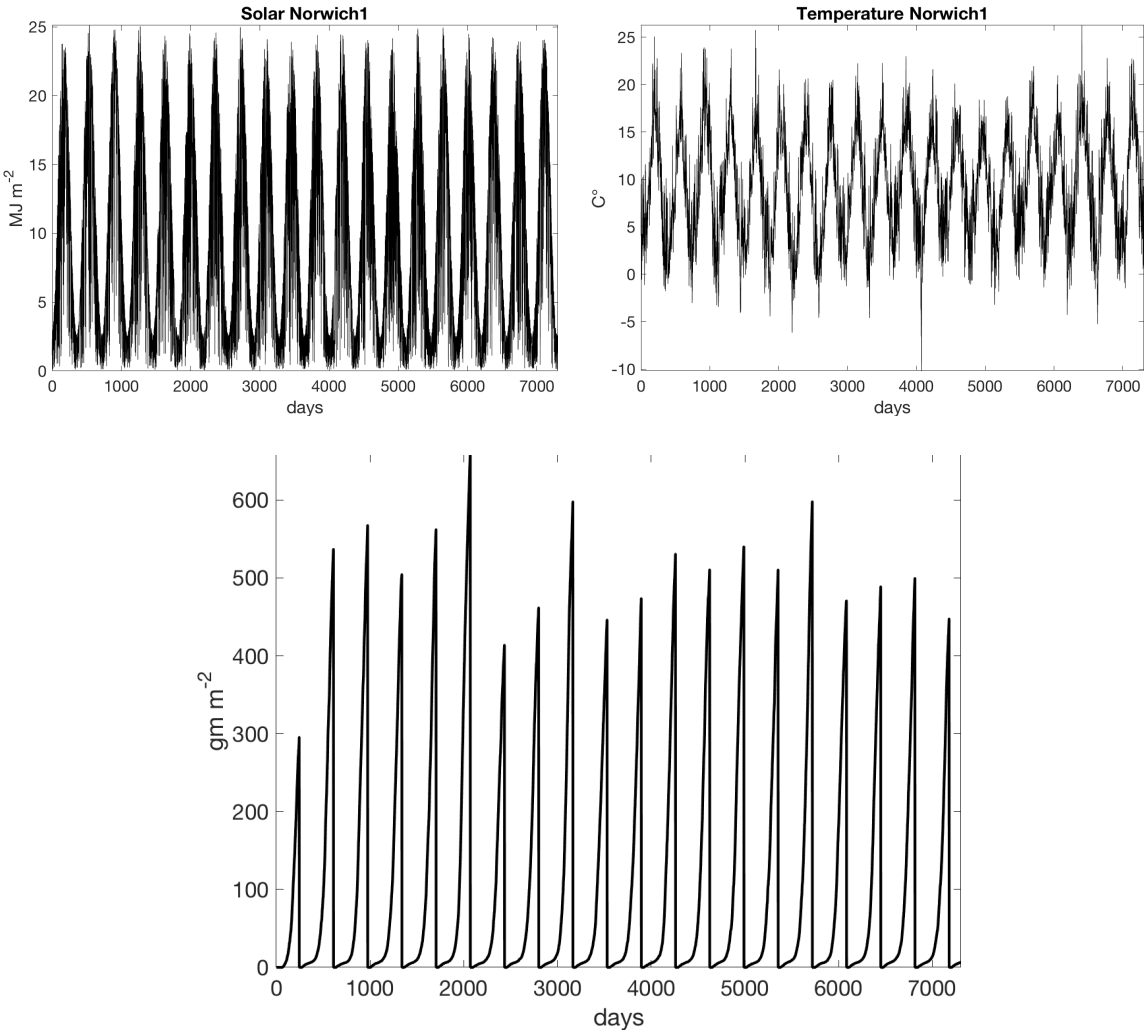

Figure 3.8. Solar radiation and temperature inputs $u_{1,1}$ (top), and crop growth model output y_0 (bottom) with θ_0 for annual grass. Simulated without solar interception by trees.

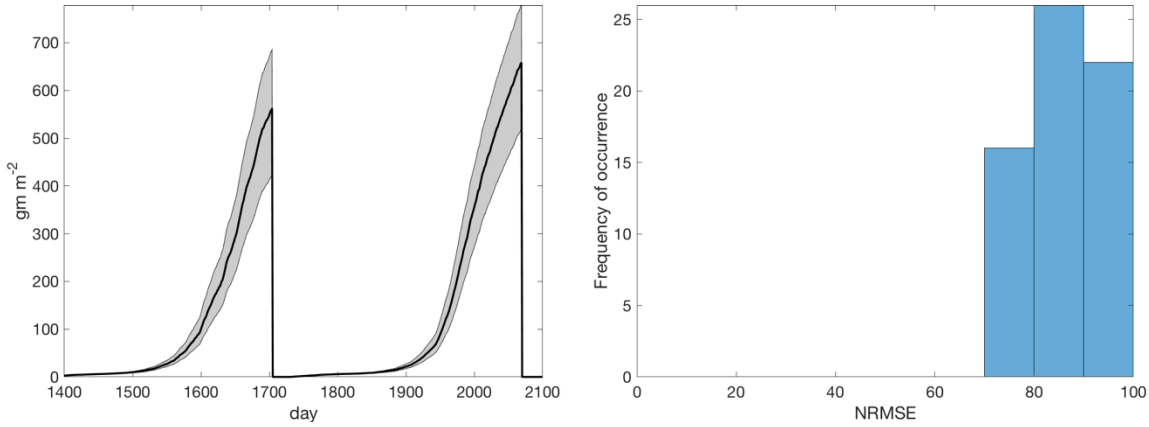

Figure 3.9. The range of annual grass outputs (days 1400–2100) for the 64 *a priori* parameter sets shown in the grey region, with the reference output in black (left) and histogram of NRMSE's for the 64 y_j (right). Model is driven by $u_{1,1}$ inputs.

3.1.4.1.3 Combination tree and crop growth model

The combination model has two outputs, one for the tree submodel and one for the crop submodel. The tree output is not affected by the crop output, while the crop output is affected by the tree output through the connecting light interception term $f_{St}(n)$. The effect of shading by the trees is clearly seen in the crop growth tapering off over time (Figure 3.10), while the tree growth is identical to the tree growth when run independently of the crop model (there are two-way effects when water and other dynamics are added).

Looking at the model error for the *a priori* best-guess parameter estimates, the combination model has 12 parameters or 2^{12} (4096) $\pm 5\%$ variations for analysis. The goodness-of-fit for the tree model is the same as shown in Figure 3.7, but the fit for the crop model is worsened by the error in the tree model through the light interception term. For the combination model, the fit for the crop falls below 90% for over 70% of the θ_j (Figure 3.11).

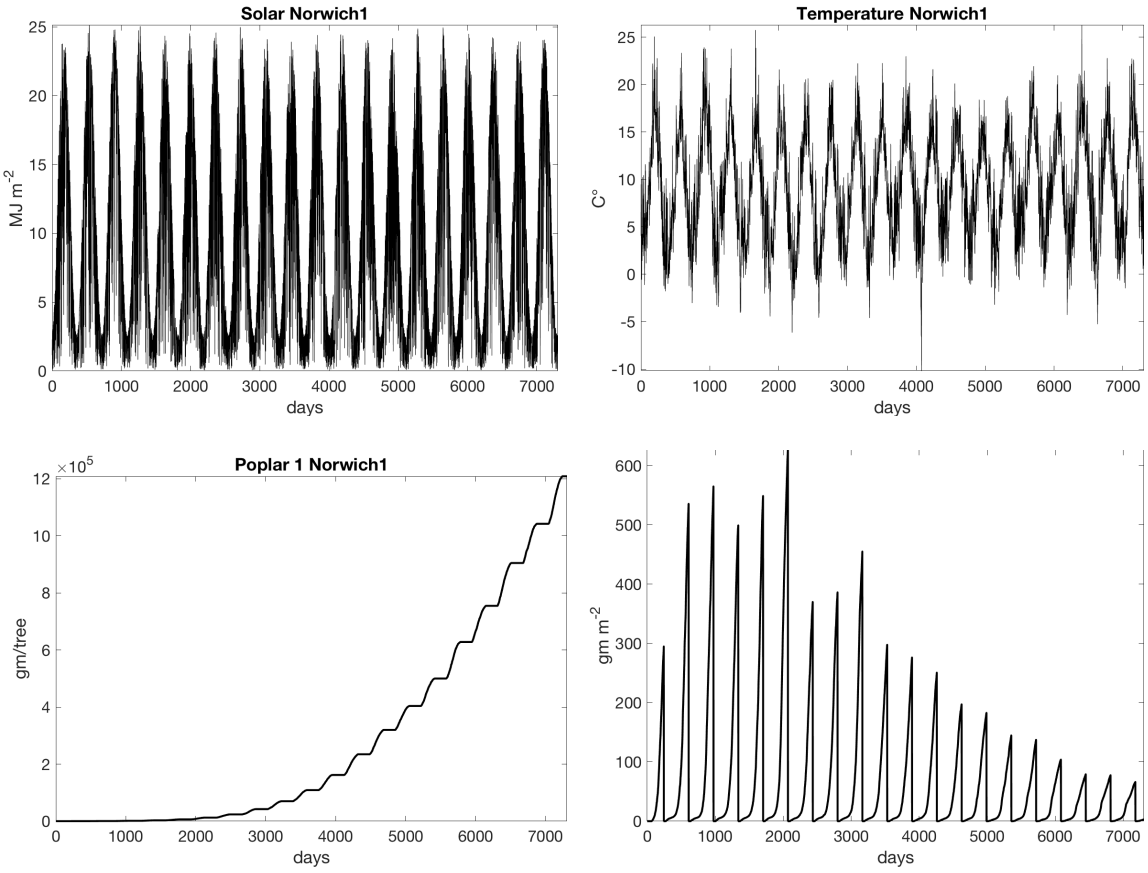

Figure 3.10. Solar radiation and temperature inputs $u_{1,1}$ (top), and combination model outputs y_0 for tree growth (bottom left) and crop growth (bottom right) with θ_0 for Poplar1 and annual grass.

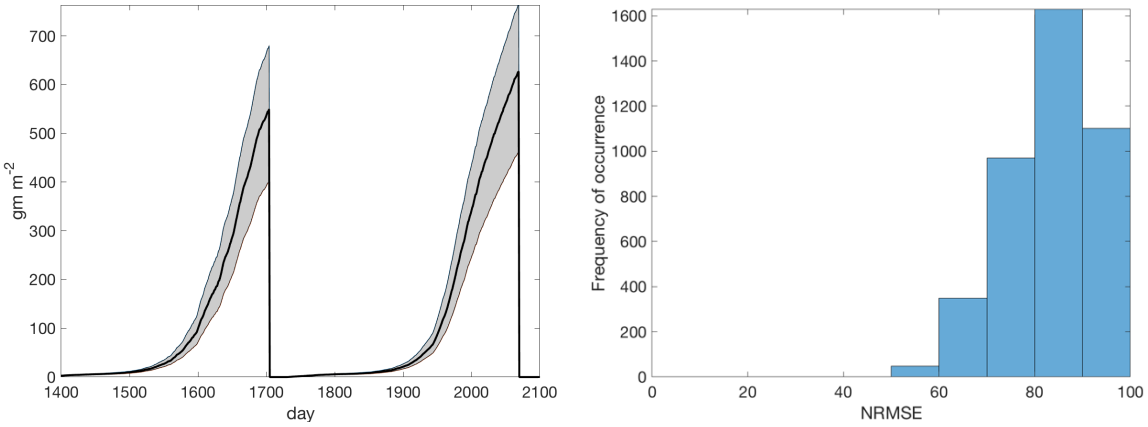

Figure 3.11. The range of annual grass outputs (days 1400–2100) for the 4096 *a priori* parameter sets shown in the grey region, with the reference output in black (left) and histogram of NRMSE's for the 4096 y_j (right). Model is driven by $u_{1,1}$ inputs.

3.1.4.2 *Perfect identification of tree growth model parameters*

Yield-SAFE was developed as "a very parameter sparse, yet process-based model" providing "the best chance that robust parameter values can be identified" (van der Werf et al. 2007). This modeling philosophy invites the question: If one had sufficient input-output (I/O) data, would one be able to use that data to identify the system parameters? This initial exploration attempts to identify the system parameters using only noise-free input-output sequences from the model. This matching system-to-system identification scheme is termed here 'perfect identification.' Although the perfect identification exercise is not part of the model reduction procedure, it can illuminate problems with parameter identifiability in practice. The perfect identification set up assesses the ability to identify the parameters in a specific model structure solely from the system's I/O data and knowledge of the model structure (Figure 3.12).

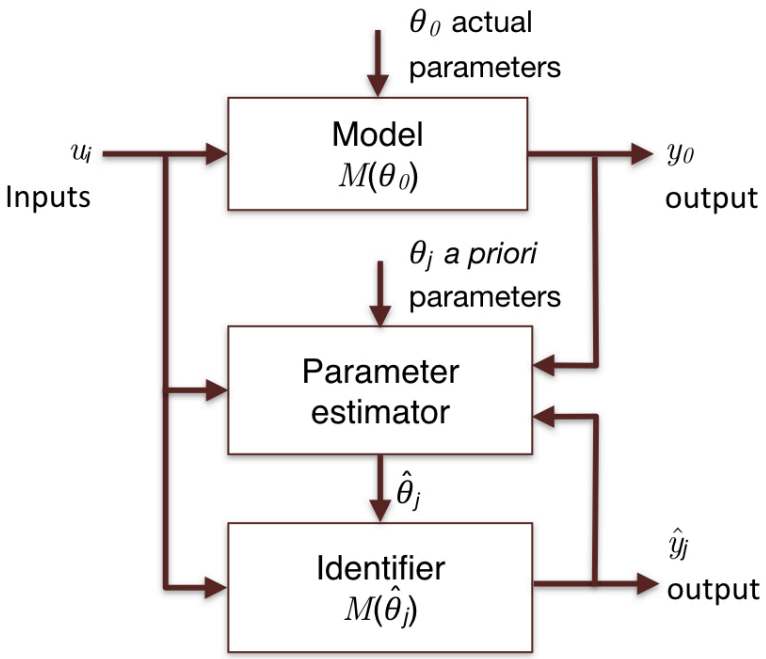

Figure 3.12: Perfect identification utilizes I/O sequences from a model to estimate model parameters θ in the identical model (Identifier) using the same I/O sequences.

Assessment in the perfect identification case consists of comparing the system parameters with the estimated parameters derived through a parameter optimization procedure. If the optimization is working for the initial conditions tested, the estimated parameters will be identical to the system parameters. Steps to perform parameter identification in the perfect identification set-up include:

1. Implement model on a computational platform including all recommended equations, variables, parameters, initial values, and constraints.
2. Run the full model with test inputs for a duration suitable to represent all long and short-term dynamics included in the model. The appropriate time scales of interest should be known from the input class and the submodel dynamics.
3. Evaluate ability to identify parameters based on the input-output sequences generated in Step 2. Extensive work has been done in numerical parameter identification for nonlinear systems (Sjöberg et al. 1995; Ljung 1983, 1999, 2010). The identification process requires
    - Input-output data collected with an appropriate sampling period (done in Step 2).
    - A performance measure or cost function to be minimized through optimization.
    - A set of initial parameter estimates.
    - A parameter optimization procedure.
    - S method for assessing accuracy.

If parameters can be identified from initial estimates within the feasible parameter space and for all test inputs, then one can conclude that the model is identifiable in practice. Determining conditions for successful perfect

identification of model parameters such as on input excitation, optimization method, and parameter initialization are all important considerations for parameter identification (Ljung 1999). However, the goal is not to find the best parameter initialization or optimization scheme, but to determine if reasonable optimization efforts lead to parameter identifiability in the perfect case.

In cases where the estimated parameters do not match the system parameters, there may be several causes

1. Parameters may be functionally related (or nearly so from a numerical standpoint) in ways that are not apparent from the model (Hengl et al. 2007).
2. The inputs may not be sufficiently exciting to all system modes, rendering certain parameters unimportant in determining system outputs. This can be due to model modes that have little impact on the output or overparameterization (Young 2013), which can cause the estimated parameters to be faulty in a predictor for an input that was not used for calibration (Whittaker et al. 2010).
3. The performance measure of the optimization procedure may be a poor choice for the system. For example, should performance measures emphasize matching the time history of outputs or its increments?
4. Initial parameter estimates may not lie in regions of attraction for the given optimization scheme and performance measure.

3.1.4.3 Parameter estimation simulations

The purpose of the present simulations is to illustrate potential challenges of parameter identification based only on knowledge of input and output data in an already parameter-sparse process-based model. In perfect identification, starting

from an initial parameter location near the correct parameters, estimation leads to identification of the correct parameters. As will be shown for this model, a conventional parameter estimation method converges to a manifold in the parameter estimate space where the model output $\hat{y}$ is nearly identical to the reference system output y_0, but where the parameter estimate $\hat{\boldsymbol{\theta}}_j$ vector can be far from the true parameters $\boldsymbol{\theta}_0$.

The tree growth model based on (3.1)–(3.5) was implemented in the MATLAB® System Identification Toolbox™ to take advantage of advanced optimization routines and the ability to easily run multiple simulations. Actual parameters $\boldsymbol{\theta}_0$ for Poplar1 (Table 3.1) and input $u_{1,1}$ (Table 3.5) were used for these simulations. MATLAB's `nlgrey` model object defined the model equations and the `nlgreyest` parameter estimation function was used with the default optimization settings. The default optimization algorithm used is the Trust-Region Reflective Newton method for nonlinear least-squares (Ljung 2018). For more details on the simulation implementation in MATLAB, see Chapter 4. The residuals and cost function are the same as those defined in Chapter 2.

The optimization was initialized with parameter estimates randomly selected on a hypersphere of normalized radius 0.5 centered on the correct parameter location (using MATLAB's `randsphere` function). From 272 initial estimates on the hypersphere, the resulting optimized parameters resided in locations in the 6-dimensional parameter space where the cost function was nearly 0 ($<10^{-5}$ gm^2 tree^{-2}). These 272 optimized parameter settings can be seen as lying on a manifold where the model output nearly matches that of the true parameters for

the given input sequence. Table 3.8 compares the optimized parameters to the true parameters in the percentage error space.

Table 3.8. Parameter error percentage ranges for 272 sets of optimized parameters determined from initial estimates randomly selected on a hypersphere around the correct parameters. All sets of estimated parameters result in a very small cost function ($<10^{-5}$ gm^2 tree^{-2}) over the 20-year simulation, i.e., the output sequences are very nearly identical for all 272 parameterizations.

Parameter	Min % error	Max % error	Mean abs % error
ε_t	-0.00	0.00	0.00
α	-0.00	0.00	0.00
L_m	-0.00	0.00	0.00
$B_t(0)$	-0.00	0.00	0.00
N(0)	-30.30	50.94	6.77
R	-51.42	127.82	13.71

Results from this experiment demonstrate that for the given input, ε_t, α, L_m, and $B_t(0)$ converge close to their actual values, while a manifold exists in N(0)–R parameter space that passes through their true values, but extends well beyond (Figure 3.13). In other words, any N(0)–R parameterization on this manifold gives near perfect fit to the true system output. This result shows that for consistent identifiability certain combinations of parameters should not be simultaneously estimated. Although these parameter estimation simulations do not prove that perfect identification would not be possible for any input signal, parameterization, or optimization scheme, they do suggest that in practice perfect identification of the full system is at best problematic. Dupraz et al. (2005, p. 99) gives another example of a manifold encountered in the Yield-SAFE parameter space with small error (between ε_t and γ, a water module parameter).

Furthermore, the results confirm that a reduced-order parameter space is advisable (or even necessary in this example) for parameter identifiability. One

could choose to reformulate the model to eliminate detected identifiability problems. However, in this case where the model is believed to give a good quality of fit to real data and is parsimonious, we follow the lead of Yield-SAFE's developers: locate a subset of the parameter space for optimization (Dupraz et al. 2005, van der Werf et al. 2007, Graves et al. 2010, Palma et al. 2017). More specifically, the procedure seeks to find a minimal number of identifiable parameters to optimize for an acceptable fit to the reference I/O data.

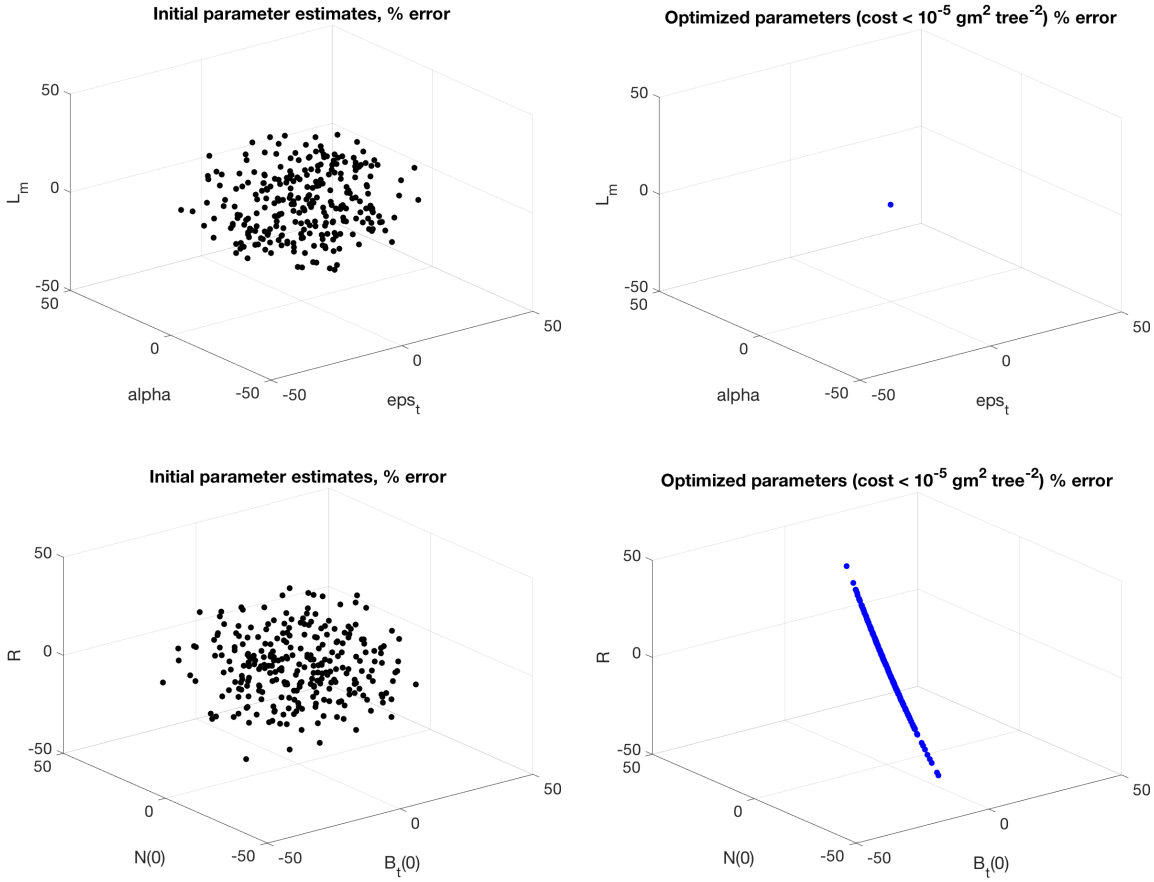

Figure 3.13: Three-dimensional plots showing locations of 272 *a priori* parameter settings on a hypersphere about the true parameters (left) and their locations after optimization (right) in the percent parameter error space. All $\hat{\theta}$ give a cost of $<10^{-5}$ gm^2 tree^{-2}. ε_t, α, L_m, and $B_t(0)$ converge to their actual values, while N(0) and R values do not. This test of identifiability reveals the presence of a manifold in the N(0)–R parameter space that extends out a considerable distance from the actual parameter values, where the cost function is very nearly zero for the given input.

3.1.4.4 One-at-a-time parameter sensitivity

A one-at-a-time (OAT) parameter sensitivity analysis is commonly used in modelling to shed light on how changes in each parameter independently impact model output (Saltelli and Annoni 2010). Van der Werf et al. (2007) state regarding Yield-SAFE, "A sensitivity analysis is presented to elucidate which biological parameters most influence short- and long-term productivity and land equivalent ratio." In describing a method to parameterize Yield-SAFE, Keesman et al. (2011) uses a normalized parameter sensitivity analysis to discern which parameters are associated with dominant model processes and therefore should be prioritized as candidates for estimation. Their analysis compared the model output at a certain time point with each parameter set to $\pm 10\%$ of its true (actual) value (with all other parameters held to their true values).

It is widely noted that OAT sensitivity analysis has serious shortcomings in the investigation of complex nonlinear systems (Saltelli and Annoni 2010). First, the method is usually applied to a single point (often the actual or nominal setting) in the parameter space. Hornberger and Spear (1981) overcome this by evaluating sensitivity throughout the parameter space. A second shortfall of OAT analysis is the failure to consider important interactions between parameters that occur in nonlinear models. Dupraz et al. (2005) noted that parameter interactions in Yield-SAFE may be important, stating, "... interactions between [the] dominant parameters can be further analysed, but for the time being we focus on [OAT] main effects only." Keesman et al. (2010) noted the limitations of OAT analysis citing more advanced sensitivity analysis of Abussam et al. (2001) and Ziehn and Tomlin (2009), but proceeded with an approach based on OAT analysis for Yield-SAFE.

Despite acknowledged shortcomings, Yield-SAFE's lead is followed here with the question: Does OAT help determine which parameters are most important to estimate for achieving a good model fit? In the OAT analysis presented here, two assumptions are made that are different from those of Keesman et al. (2010). First, the reference output y_0 is generated using the actual parameters, without assuming that the actual parameters are known by the modeler. Instead, the OAT analysis is conducted in the neighborhood of an *a priori* parameter location where all parameters are $\pm 5\%$ off of their true values. Specifically, we use the *a priori* parameter estimate θ_1, in which all parameters are –5% away from their actual values (Table 3.9). Second, instead of comparing y_j at one point in time, the NRMSE goodness-of-fit measure is employed comparing $y_0(n)$ with $y_j(n)$ for the entire time series.

Table 3.9 Actual model parameters (Poplar1) and a priori settings for OAT sensitivity analysis. The *a priori* parameters are all –5% away from their actual values. NRMSE values are calculated at 100 equidistant points within ±20% of each of the *a priori* parameters.

Parameter	θ_0 Poplar1	θ_1 Poplar1
ε_t	1.4086	1.33817
α	0.0001	0.000095
L_m	500	475
$B_t(0)$	100	95
N(0)	0.6225	0.591375
R	200000	190000
τ	10	10
k_t	0.8	0.8
k_a	10	10
k_b	0.4	0.4

The results of this analysis are shown in Figure 3.14, which leads to several observations. First, adjustment of any one of 5 of the 6 parameters (not α) can

increase NRMSE from about 79% to above 95%, which might be considered by the modeler to be a good enough fit. It is unclear which of these 5 parameters should be adjusted, as adjusting any one can give a much improved fit. Second, and perhaps more important, it would be impossible to discern from Figure 3.14 that the true parameters lie at about +5.26% of the user's estimated values. Namely, three of the six parameters need to move in a direction that *worsens* NRMSE values from an OAT perspective. This observation suggests that OAT analysis may help select parameters for optimization, but that it is of little help in locating the actual parameters. This also indicates that when selecting a subset of parameters to estimate via OAT analysis, the optimized parameter values may lose their biophysical interpretation. In other words, optimizing the parameters that are most sensitive from a OAT perspective, may drive parameters away from their actual values (assuming that the full model is an accurate version of the underlying biophysical processes). This observation based on Figure 3.14 may explain why Dupraz et al. (2005, p. 98) observed an occasion where estimating the most important parameter from the water module (γ) and fixing another (ε) from an OAT perspective led to an optimized parameter solution outside of the *a priori* feasible range. This issue is not unique to OAT analysis, but is an outcome of estimating within a reduced-order parameter space.

A similar simulation with the *a priori* parameter estimate $\boldsymbol{\theta}_{64}$, in which all parameters are +5% away from their actual values, also suggests that OAT analysis is difficult to interpret for selecting which parameters are most or least important for estimation (Figure 3.15). Another way to view this observation is that the interactions between parameters have complex and perhaps nonintuitive effects on model fit as measured by NRMSE. This points towards the advantage

of looking at the Hessian of the cost function to determine the most important directions in the parameter space for minimization (Keesman 1989).

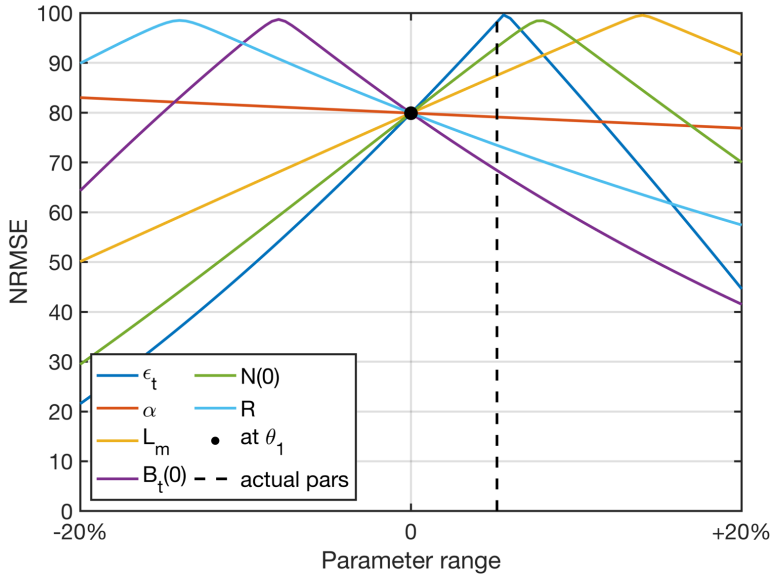

Figure 3.14 One-at-a-time parameter sensitivity in the neighborhood (±20%) of the *a priori* setting θ_1 (all parameters –5% away from their actual values) in terms of goodness-of-fit measure NRMSE (perfect fit = 100). NRMSE values are calculated at 100 equidistant points within ±20% of each of the *a priori* parameters. The output is driven by the $u_{1,1}$ input. All actual parameters lie at the dashed line at +5.26% of their *a priori* values.

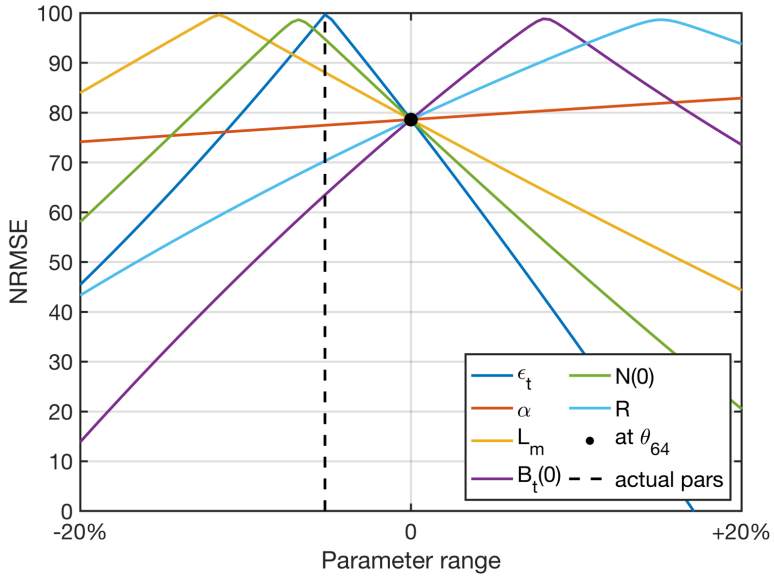

Figure 3.15 One-at-a-time parameter sensitivity in the neighborhood (±20%) of the *a priori* setting θ_{64} (all parameters +5% away from their actual values) in terms of goodness-of-fit measure NRMSE (perfect fit = 100). The output is driven by the $u_{1,1}$ input. All actual parameters lie at the dashed line at -4.76% of their *a priori* values.

3.2 HIROPE procedure applied to Yield-SAFE tree growth model

The following example for the Yield-SAFE model illustrates the procedure described in Section 2.7 for reducing the parameter space for optimization. As noted above, in this analysis, solar radiation and temperature are considered limiting (as well as being the only growth drivers), while water and soil nutrients will be considered non-limiting, effectively removing the associated water and nutrient dynamics for initial study of the model reduction procedure.

For the following analysis of the tree model, the following are used

Parameter vector $\Theta = [\ \varepsilon_t\ \alpha\ L_m\ B_t(0)\ N(0)\ R\ ]$

Actual parameters θ_0 = Poplar1 (from Table 3.2)

A priori ±5% adjustments in parameters θ_j are given in Table 3.7

A list of procedure steps and expected outcomes are given in Table 3.10.

Table 3.10. Simulation list for tree model. Action steps are repeated from the HIROPE procedure of Section 2.7.

M_t (tree model)	
Action	Expected outcome
Step 1. Compare goodness-of-fit of y_0 with the estimated output y_j. For one u_i, compare output y_0 for a parameter set θ_0 with the output y_j for one of the associated *a priori* θ_j parameter set using the NRMSE criterion.	It is expected that the NRMSE value will be less than the cut-off measure of 99%, meaning that parameter optimization must be done on θ_j for the model to have acceptable performance.
Step 2. Compute $H_{i,j}$ and its eigenvectors and eigenvalues. Compute $H_{i,j}$ for θ_j using data sequences u_i, y_0, and y_j from above simulation and determine its eigenvectors and eigenvalues.	The condition number of $H_{i,j}$ is expected to be ≥10^6, meaning that some parameters will be associated with steep curvature of the loss function and other parameters will be associated with shallow curvature of the loss function.

Step 3. Generate Λ parameter ranking. From the computed $H_{i,j}$, rank parameters in order of highest to lowest magnitude of associated eigenvalues (projecting eigenvectors onto the axes of their largest associated parameters). In cases where the projection of an eigenvector is equal in magnitude for 2 or more parameters, choose one of them, then both in the ranking.	This process yields a ranking of reduced-order parameter sets $\chi_{j,k}$ for optimization.
Step 4. Confirm consistent identifiability. Beginning with highest ranked reduced-order parameter set $\chi_{j,1}$ containing only one parameter associated with the largest eigenvalue, optimize the parameter value from both of its ±5% values (with all other parameters in θ_j fixed to their original estimated values) to determine a $\hat{\theta}_j$.	It is expected that from both initializations, the optimized parameter value will be the same (within ±0.01%), an indication of consistent identifiability. If both optimized values are not identical, then the model is not consistently identifiable for the selected $\chi_{j,k}$. In such a case skip the current $\chi_{j,k}$, and move to the next without estimating the last added parameter in future $\chi_{j,k}$.
Step 5. Check goodness-of-fit of y_0 and $\hat{y}_j$. Using the current $\hat{\theta}_j$, generate $\hat{y}_j$ and calculate the NRMSE value relative to y_0. If criterion (NRMSE ≥99%) is met, then move to the next step for validation. If optimization on the first candidate parameter set does not meet the NRMSE criterion, then repeat Step 4 for the next candidate parameter set, and so on, until either the fitting criterion is met for the optimization and validation inputs (success) or not (failure of procedure). In case of failure, revise fit criterion or model.	It is expected that the first candidate parameter set (one parameter) will improve model performance for both the over- and understory models, but it is uncertain how many parameters are needed to meet the criterion.
Step 6. Validation. Validate the optimized parameter set using a number of different input sequences from the same location and calculating their NRMSE values. If the validating input data also results in an acceptable NRMSE value, then the model optimization procedure is considered successful, otherwise return to Step 4. One may proceed to the next candidate parameter set if one wishes to see if there is improvement.	It is expected that validation with another input from the same location will give a similar fit to the optimization input.
Simulations to show robustness	
Validate parameterizations using 4 inputs from the same location L_i.	Expected that the validation inputs will work well.

3.2.1 Prerequisites tree model

The prerequisites for the procedure applied to the tree model are as follows

- The full system equations are defined. See equations (3.1)–(3.5).
- Determine a priori conditions on parameters.

Equations (3.1)–(3.5) are twice continuously differentiable with respect to the parameters except when $B_t(n)+\frac{B_t(0)}{N(0)}\left((L_m R)^{1/2}-1\right)=0$. However, not all parameter values give realistic simulation outputs, which could become relevant in parameter estimation. For example, if α is too large, a negative biomass could result. If $(L_{max}R)^{1/2}-1<0$ then the number of shoots can become negative or if $B_t(n)+\frac{B_t(0)}{N(0)}\left((L_m R)^{1/2}-1\right)$ becomes a very small number, the number of shoots can grow very large. Table 3.2 shows that α is a small number and $(L_{max}R)^{1/2}\gg 1$ for all Yield-SAFE parameterizations documented in Burgess et al. (2014), so in practice one can consider the model to be continuous with feasible outputs over the feasible parameter range.

a. The model is implemented on a computational platform. Here the model is implemented in MATLAB 2018a.
b. If model equations are not twice continuously differentiable with respect to the parameters of interest, acceptable differentiable analytic approximations are available. All equations are twice continuously differentiable.
c. General class of model inputs is known and ample input data at the model time step can be accessed. Inputs are synthesized by climate models and accessed through CliPick (see Section 3.1.3).
d. Acceptable a priori parameter values and initial conditions are available. The

Burgess et al. (2014) parameterizations are used as the actual parameters for generating reference outputs and ±5% variations of those parameters for *a priori* test parameterizations.

e. Cost function is defined. A standard sum squared error is used.

f. Measure of fit function is defined, and a cut-off criterion for goodness-of-fit selected. NRMSE is used with a cut-off threshold of 99% for an acceptable fit.

g. A measure of consistent identifiability for parameters is defined. Here defined as optimized parameters from different *a priori* locations within 0.01% of the same location.

h. A parameter optimization routine is implemented. The Trust-Region Reflective Newton algorithm as implemented in MATLAB via the `lsqnonlin` function was used (see Section 4.1.1).

Step 1. Compare goodness-of-fit of y_0 with the estimated output y_1.

Input u_{11} with actual parameter set Poplar1 generates y_0, and *a priori* parameter set θ_1 generates the output realization y_1 (see Table 3.11 for parameter values). As seen in Figure 3.16, the *a priori* model fit is unacceptable, with a NRMSE of 79.09%.

Table 3.11 Actual parameters θ_0 *a priori* settings θ_1 for tree model

Parameter	θ_0	θ_1
ε_t	1.4086	1.33817
α	0.0001	0.000095
L_m	500	475
$B_t(0)$	100	95
N(0)	0.6225	0.591375
R	200000	190000

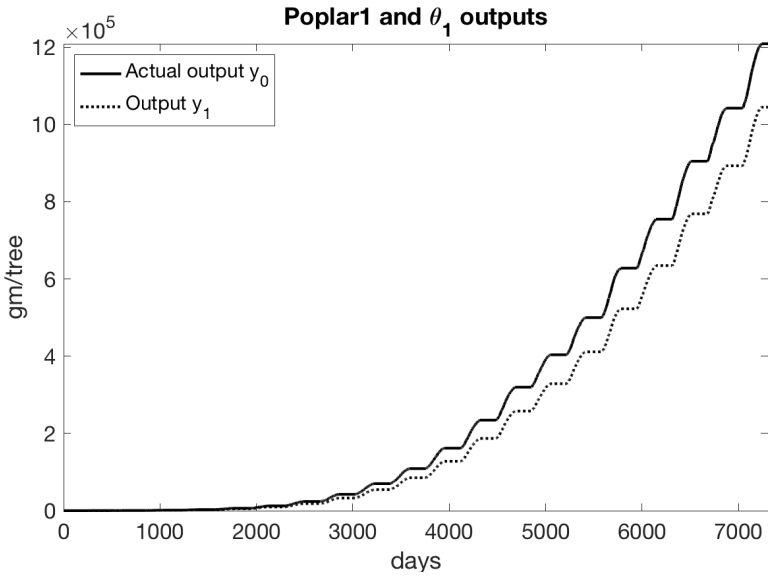

Figure 3.16 Outputs y_0 and y_1 for tree model driven by $u_{1,1}$. The NRMSE is 79.09%, indicating parameter must be adjusted to achieve an acceptable fit.

Step 2. Compute $H_{i,j}$ and its eigenvectors and eigenvalues.

An analytical expression for the Hessian of the cost function begins with expressions for the error residual and cost function

$$r(n) = y_0(n) - y_j(n) \tag{3.11}$$

$$\begin{aligned} V(n) &= \frac{1}{2}\sum_{i=1}^{n} r^2(i) \\ &= \frac{1}{2}\sum_{i=1}^{n}\left(y_0(i) - y_j(i)\right)^2 \end{aligned} \tag{3.12}$$

By combining equations (3.1)–(3.5) into one expression yields

$$\begin{aligned} y_j(n) = B_t(n) &= B_t(n-1) \\ &+ \frac{\varepsilon_t}{\rho} I(n)\left(1 - e^{-\rho\left(k_t - k_b + k_b\left(\frac{k_a}{L_t(n)+k_a}\right)\right)\left(L_t(n-1) + \left(1-e^{-\frac{1}{\tau}}\right)\left(\left(\frac{L_m}{R}\right)^{1/2} N(n-1) - L_t(n-1)\right)\right)\phi(n)}\right) \\ &- \alpha B_t(n-1) \end{aligned} \tag{3.13}$$

where from (3.2)

$$N(n-1) = N(n-2) + (L_m R)^{1/2} \left(\frac{\frac{B_t(0)}{N(0)}\left((L_m R)^{1/2} - 1\right)}{\left(B_t(n-1) + \frac{B_t(0)}{N(0)}\left((L_m R)^{1/2} - 1\right)\right)^2} \right) \times \left(B_t(n-1) - B_t(n-2) \right) \tag{3.14}$$

By (2.32) the partial derivates of the sum of the squared residuals are the same as the sum of the partial derivatives of the squared residuals. In order to implement parameter scaling, introduce new symbolic variables s_1, s_2, ... s_6 paired with each parameter as shown in Table 3.12 (see previous example for logistics equation, Section 2.6.1). The addition of the symbolic variables leads to revised expressions for the model used to derive the Hessian

$$y_j(n) = B_t(n) = B_t(n-1) + \frac{s_1 \varepsilon_t}{\rho} I(n) \left(1 - e^{-\rho\left(k_t - k_b + k_b\left(\frac{k_a}{L_t(n)+k_a}\right)\right)\left(L_t(n-1) + \left(1-e^{-\frac{1}{\tau}}\right)\left(\left(\frac{s_3 L_m}{s_6 R}\right)^{1/2} N(n-1) - L_t(n-1)\right)\right)\phi(n)} \right) - s_2 \alpha B_t(n-1) \tag{3.15}$$

with

$$N(n-1) = N(n-2) + (s_3 L_m s_6 R)^{1/2} \left(\frac{\frac{s_4 B_t(0)}{s_5 N(0)}\left((s_3 L_m s_6 R)^{1/2} - 1\right)}{\left(B_t(n-1) + \frac{s_4 B_t(0)}{s_5 N(0)}\left((s_3 L_m s_6 R)^{1/2} - 1\right)\right)^2} \right) \times \left(B_t(n-1) - B_t(n-2) \right) \tag{3.16}$$

Table 3.12 Tree model parameters paired with symbolic variables prior to differentiation for Hessian. Parameter scaling is achieved by setting the model parameters to their *a priori* values and differentiating with respect to the symbolic variables s_1, s_2, ... s_6, then setting the symbolic variable values to 1.

Parameter		Parameter with symbolic variable
ε_t	→	$s_1\varepsilon_t$
α	→	$s_2\alpha$
L_m	→	$s_3 L_m$
$B_t(0)$	→	$s_4 B_t(0)$
$N(0)$	→	$s_5 N(0)$
R	→	$s_6 R$

The Hessian of the squared residual, which can be determined analytically, is the basis for calculating the Hessian over n steps

$$H\left[r^2(i)\right] = H\left[\left(y_0(i) - y_j(i)\right)^2\right] \tag{3.17}$$

The Hessian of (3.17) with respect to the six parameters of interest ε_t, α, L_m, $B_t(0)$, $N(0)$, and R (with scaling as described above) was formulated using MATLAB's symbolic `hessian` function, yielding an 6 x 6 matrix that can be evaluated at each time i. Calculation of the Hessian followed the steps given in the analysis of the logistic equation in Section 2.6. For this analysis, the Hessian was calculated for N=7306 with eigenvalues and eigenvectors given in Table 3.13. One can see that the eigenvalues of the Hessian have a four order of magnitude range (condition number 1.9e–04). Note that the Hessian has two negative eigenvalues, indicating that the cost function surface is nonconvex.

Table 3.13 Hessian with eigenvalues and eigenvectors, ordered by eigenvalue magnitude.

	Hessian (N=7306)					
	3.49e+05	-2.54e+04	-1.34e+06	1.26e+03	-1.26e+03	1.34e+06
	-2.54e+04	2.46e+03	-6.78e+02	5.48e-01	-5.48e-01	6.77e+02
	-1.34e+06	-6.78e+02	7.31e+05	1.24e+03	-1.24e+03	6.19e+05
	1.26e+03	5.48e-01	1.24e+03	-1.24e+03	-2.78e+01	5.11e+01
	-1.26e+03	-5.48e-01	-1.24e+03	-2.78e+01	1.29e+03	-5.11e+01
	1.34e+06	6.77e+02	6.19e+05	5.11e+01	-5.11e+01	-1.97e+06
	Eigenvalue					
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
	-2.9463e+06	1.9472e+06	1.1226e+05	1.3230e+03	-1.2763e+03	5.6379e+02
	Eigenvector					
Axis	X_1	X_2	X_3	X_4	X_5	X_6
ε_t	-0.4635	0.6980	-0.5409	-0.0114	0.0142	-0.0713
R	0.8305	0.1274	-0.5372	-0.0110	0.0146	-0.0713
L_m	-0.3089	-0.7046	-0.6345	-0.0101	0.0155	-0.0715
α	-0.0043	-0.0088	0.1259	-0.2525	0.0971	-0.9544
N(0)	-0.0003	-0.0000	0.0135	0.9672	0.0040	-0.2537
$B_t(0)$	0.0003	0.0000	-0.0132	-0.0213	-0.9949	-0.0974

Step 3. Generate Λ parameter ranking.

Based on Table 3.13, the Λ values and their associated ranking for this *a priori* setting are shown in Table 3.14.

Table 3.14 Λ values and their associated ranking based upon Hessian eigenstructure.

Ranking	**Parameter**	Λ
1	ε_t	2.78e+06
2	R	2.76e+06
3	L_m	2.35e+06
4	α	4.49e+04
5	N(0)	3.87e+03
6	$B_t(0)$	3.76e+03

This leads to the following $\chi_{1,j}$ candidate reduced-order parameter vectors for estimation (Table 3.15).

Table 3.15 $\chi_{1,j}$ reduced-order parameter space candidates for tree model.

	Ranking					
	$\chi_{1,1}$	$\chi_{1,2}$	$\chi_{1,3}$	$\chi_{1,4}$	$\chi_{1,5}$	$\chi_{1,6}$
ε_t	X	X	X	X	X	X
R		X	X	X	X	X
L_m			X	X	X	X
α				X	X	X
N(0)					X	X
$B_t(0)$						X

Step 4. Confirm consistent identifiability.

$\chi_{1,1}$ designates that ε_t is the first parameter to estimate. Optimizations from both *a priori* settings at $\pm 5\%$ of the actual ε_t value (θ_1 and θ_2) converge to the same estimate, as shown in Table 3.16, confirming consistent identifiability.

Table 3.16 Actual parameters and ±5% *a priori* settings θ_1 and θ_2 for the first ranked candidate parameter R together with the optimized setting. For comparison, θ_0 is also included.

Parameter	θ_1	θ_2	$\hat{\theta}_1$ and $\hat{\theta}_2$	θ_0
ε_t	1.3382	1.4790	1.4145	1.4086
R	190000	190000	190000	200000
L_m	475	475	475	500
α	0.000095	0.000095	0.000095	0.0001
N(0)	0.591375	0.591375	0.591375	0.6225
$B_t(0)$	95	95	95	100

Step 5. Check goodness-of-fit of y_0 and $\hat{y}_j$.

The NRMSE value for the model with θ_1 and ε_t estimated is 99.92%, which meets the criterion of ≥99%.

Step 6. Validation.

The next step is validating the model behavior using $\hat{\theta}_1$ for additional inputs from the same location L_1 (Table 3.6). For this test, the model with $\hat{\theta}_0$ is run with the validation inputs to generate new reference outputs, which are compared with the model run with the validation inputs with $\hat{\theta}_1$. As shown in Table 3.17, the model fit with $\hat{\theta}_1$ is similar for all validation sequences, suggesting that the estimated parameterization has met all model criteria. Therefore, this optimized setting is now called $\hat{\hat{\theta}}_1$.

Table 3.17 Fit measures for $\hat{\hat{\theta}}_1$ with ε_t estimated.

Input	Use	NRMSE	Cost gm^2 tree^{-2}
$u_{1,1}$	Optimization	99.92%	7.5473e+04
$u_{1,2}$	Validation	99.92%	1.2323e+05
$u_{1,3}$	Validation	99.91%	1.2090e+05
$u_{1,4}$	Validation	99.93%	7.9129e+04
$u_{1,5}$	Validation	99.93%	7.0020e+04

Even though estimation of ε_t led to an acceptable model, can the estimate be improved by returning to Step 4 and the next ranked subset of parameters, $\chi_{1,2}$? To answer this question, we return to Step 4 with the next reduced-order parameter set.

Step 4 (Extra iteration on $\chi_{1,2}$). Confirm consistent identifiability.

In $\chi_{1,2}$, both ε_t and R are estimated. As an indicator of identifiability, the optimization is carried out from all ±5% combinations of each parameter or *a priori* parameter sets θ_1, θ_2, θ_{33}, and θ_{34} (Table 3.18).

Table 3.18 Actual parameters and ±5% *a priori* settings θ_1 and θ_{33} for the first ranked candidate parameter R together with the optimized setting. For comparison, θ_0 is also included.

Parameter	θ_1	θ_2	θ_{33}	θ_{34}	$\hat{\theta}_1,\hat{\theta}_2,\hat{\theta}_{33},\hat{\theta}_{34}$	θ_0
ε_t	1.33817	1.4790	1.33817	1.4790	1.4131	1.4086
R	190000	190000	210000	210000	189340	200000
L_m	475	475	475	475	475	500
α	0.000095	0.000095	0.000095	0.000095	0.000095	0.0001
N(0)	0.591375	0.591375	0.591375	0.591375	0.591375	0.6225
$B_t(0)$	95	95	95	95	95	100

Step 5. Check goodness-of-fit.

The NRMSE value for the model for $\hat{\theta}_1$ with ε_t and R estimated is 99.96%, slightly better than that with only ε_t estimated.

Step 6. Validation.

Once again the model is run with validation sequences to confirm the parameter estimates are robust. Of note here, the fit is only slightly better than that achieved when only ε_t was estimated, and for some validation sequences the fit is slightly worse (Table 3.19).

Table 3.19 Fit measures for $\hat{\hat{\theta}}_1$ with ε_t and R estimated as compared with ε_t only estimated.

		Estimate ε_t		Estimate ε_t and R	
Input	**Use**	**NRMSE**	**Cost gm^2 tree^{-2}**	**NRMSE**	**Cost gm^2 tree^{-2}**
$u_{1,1}$	Optimization	99.94%	4.7253e+04	99.96%	3.9903e+04
$u_{1,2}$	Validation	99.94%	4.1687e+04	99.95%	4.6079e+04
$u_{1,3}$	Validation	99.92%	9.9951e+04	99.96%	3.8313e+04
$u_{1,4}$	Validation	99.94%	5.0457e+04	99.91%	1.2603e+05
$u_{1,5}$	Validation	99.94%	4.4677e+04	99.89%	1.5023e+05

3.2.2 Further analysis of procedure applied to tree model

3.2.2.1 Cost function contours at θ_1

That estimation of the two parameters ε_t and R yields an acceptable model parameterization from θ_1, allows convenient graphical examination of the cost function in the 2-dimensional ε_t–R parameter space. A plot of the cost function contours (Figure 3.17) with ε_t and R varied by $\pm 35\%$ shows that the cost function surface appears smooth, with only one local minimum. This observation is confirmed by optimizing ε_t and R from all 8 combinations of $\pm 10\%$, $\pm 50\%$, and $\pm 90\%$ of their *a priori* values, for which all converge to the same $\hat{\hat{\theta}}_1$. This test confirms the simpler consistent identifiability test conducted above as part of the procedure.

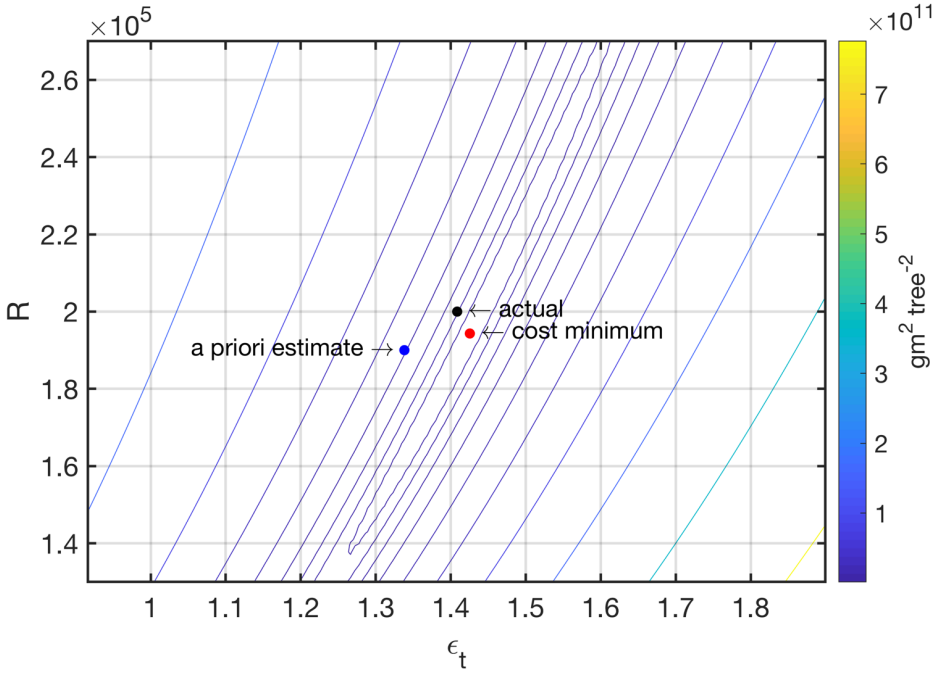

Figure 3.17 Contours of cost function for parameters ε_t and R with the other parameters fixed at their θ_0 values.

3.2.2.2 Eigenstructure at θ_0

The first question is if the eigenstructure of the cost function Hessian changes at the correct parameter setting θ_0. In other words, are the same parameters associated with the strongest curvature of the cost function at the actual parameters θ_0?

The answer to this is given by comparison of the eigenstructure at the actual parameter setting in Table 3.20 as compared with Table 3.13. The ranking of the parameters as well as the direction of the eigenvectors is different at θ_0 as compared with θ_1. At the actual parameters θ_0, the parameters rank (from largest associated eigenvalue to smallest) as ε_t, α, R, L_m, $N(0)$, $B_t(0)$, while at θ_1 the ranking is ε_t, R, L_m, $N(0)$, $B_t(0)$, α. Differences in ranking can be explained by the nature of the Hessian—for nonlinear models the Hessian is a function of the parameter estimates, meaning that it varies throughout the parameter space.

As a follow-up question, one might ask if estimates of R and L_m would converge to their actual values if the other parameters are located at their actual values. Optimization of R and L_m from all combinations of *a priori* settings $\pm 10\%$, $\pm 50\%$, and $\pm 90\%$ of their actual values all converge to their actual values. This result is further evidence that local minima do not exist in the region about the parameter origin.

Table 3.20 Hessian, eigenvalues, and eigenvectors for tree model with $u_{1,1}$, computed with parameters set to θ_0=Poplar1 for both reference and model (n=7306).

	Hessian (N=7306)					
	1.56e+05	-1.28e+04	3.80e+03	-3.40e+00	3.40e+00	-3.79e+03
	-1.28e+04	1.38e+03	-3.13e+02	1.83e-01	-1.83e-01	3.13e+02
	3.80e+03	-3.13e+02	2.11e+02	6.31e-02	-6.31e-02	-1.76e+01
	-3.40e+00	1.83e-01	6.31e-02	-1.70e-01	-8.11e-03	1.23e-01
	3.40e+00	-1.83e-01	-6.31e-02	-8.11e-03	1.87e-01	-1.23e-01
	-3.79e+03	3.13e+02	-1.76e+01	1.23e-01	-1.23e-01	-1.76e+02
	Eigenvalue					
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
	1.57e+05	3.24e+02	-2.82e+02	1.32e+02	1.87e-01	-1.71e-01
	Eigenvector					
Parameter	X_1	X_2	X_3	X_4	X_5	X_6
ε_t	-0.9960	-0.0823	-0.0282	0.0194	-0.0001	-0.0001
R	0.0242	0.0001	-0.9826	-0.1839	0.0002	0.0002
L_m	-0.0243	-0.0006	0.1833	-0.9827	0.0012	0.0011
α	0.0822	-0.9966	0.0021	-0.0010	-0.0003	-0.0003
N(0)	-0.0000	-0.0003	-0.0000	0.0011	0.9998	-0.0219
$B_t(0)$	0.0000	0.0003	0.0000	-0.0011	-0.0219	-0.9998

3.2.2.3 Simulations to show robustness

For the tree model, validation of $\hat{\hat{\theta}}_1$ (optimizing ε_t or both ε_t and R) done in Section 3.2.1 with inputs from the same location as the input used for optimization, showed very good model fits (NRMSE $> 99.9\%$). One might ask how the $\hat{\hat{\theta}}_1$ parameterization performs for inputs from other locations. To answer this question, the model was run with inputs from locations L_2–L_{10} with parameter settings both at θ_0 and $\hat{\hat{\theta}}_1$. Fit measures are given in Table 3.21. The fit is acceptable for all locations considered, with all NRMSE values $>99.7\%$. This may be somewhat surprising, as one may think that growth parameters would differ in different locations. However, it appears that in this case (non-water limiting, parameters Poplar1), that the procedure leads to parameters that

are applicable to many different locations for the tree growth model. Alternatively, one might investigate if the model itself is too insensitive to location.

Table 3.21 Fit measures for $\hat{\hat{\theta}}_1$ with ε_t and R estimated using inputs from locations other than the location used for parameter estimation.

		Estimate ε_t		Estimate ε_t and R	
Input	**Use**	**NRMSE**	**Cost gm^2 tree^{-2}**	**NRMSE**	**Cost gm^2 tree^{-2}**
$u_{1,1}$	Optimization	99.94%	4.7253e+04	99.96%	3.9903e+04
$u_{2,1}$	Validation	99.92%	7.7158e+04	99.88%	1.7693e+05
$u_{3,1}$	Validation	99.75%	8.7143e+06	99.73%	9.7128e+06
$u_{4,1}$	Validation	99.84%	1.1356e+06	99.87%	7.2095e+05
$u_{5,1}$	Validation	99.75%	7.4014e+06	99.74%	8.0533e+06
$u_{6,1}$	Validation	99.88%	4.3559e+05	99.93%	1.4147e+05
$u_{7,1}$	Validation	99.91%	1.7862e+05	99.96%	3.9592e+04
$u_{8,1}$	Validation	99.86%	7.4935e+05	99.90%	3.6380e+05
$u_{9,1}$	Validation	99.80%	2.4518e+06	99.82%	2.0398e+06
$u_{10,1}$	Validation	99.81%	1.8441e+06	99.84%	1.4112e+06

3.2.2.4 Ranking by Λ for all 64 θ_j

As a test of robustness of the Λ ranking, the Λ_i for all parameters and all 64 θ_j are calculated. As shown in Table 3.22, the relative ranking of the top four parameters is the same at all θ_j. The consistency of rank suggests some robustness to the a priori estimate within 5% of the actual parameters. Figure 3.18 gives a graphical representation of Λ_i for all θ_j.

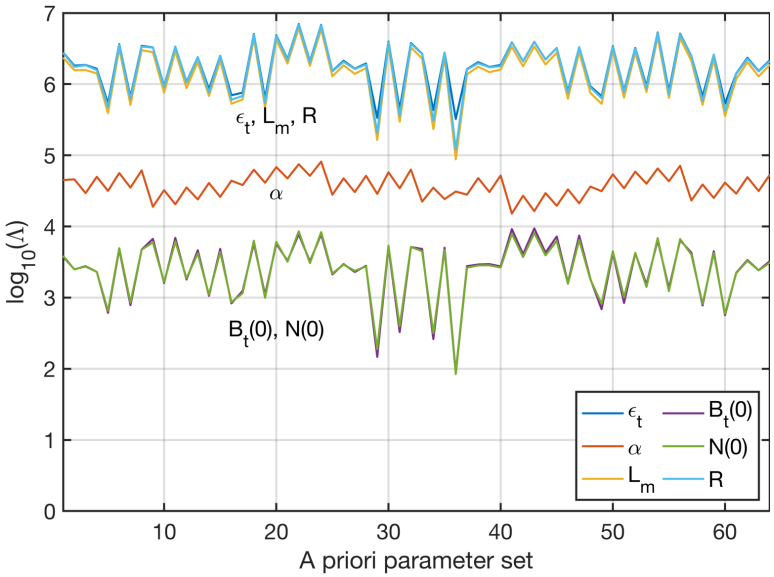

Figure 3.18 Λ_i for all parameters and all 64 θ_j in $\log_{10}$ scale. the ranking of the top four parameters is the same throughout.

Table 3.22 Occurrence of parameter ranking for all 64 *a priori* estimates θ_j.

	Ranking					
Par.	**1**	**2**	**3**	**4**	**5**	**6**
ε_t	64	0	0	0	0	0
R	0	64	0	0	0	0
L_m	0	0	64	0	0	0
α	0	0	0	64	0	0
N(0)	0	0	0	0	29	35
$B_t(0)$	0	0	0	0	35	29

3.3 HIROPE procedure applied to Yield-SAFE Understory

Reduction of the Yield-SAFE crop model follows the overstory analysis (Table 3.23). The primary difference is that the understory model description requires substitution of two model equations with continuously differentiable approximations, as detailed below.

For the following analysis of the crop model, the following notation is used:

Parameter vector $\boldsymbol{\Theta} = [\ \varepsilon_c\ \ T_0\ \ P_0\ \ S_1\ \ S_2\ \ k_c\ ]$

Actual parameters $\boldsymbol{\theta}_0$ = Annual grass (from Table 3.4)

A priori parameters $\boldsymbol{\theta}_j$ follow the example given in Table 3.7

Table 3.23. Simulation list for crop model.

M_c (crop model)	
Action	Expected outcome
Run all simulations described above for M_t.	Because the crop model is very different from the tree model, including two discontinuous functions, it is unclear how many parameters will need to be estimated to achieve a good fit.

3.3.1 Prerequisites

The procedure has the following prerequisites

a. The full system equations are defined. Given in (3.6)–(3.10). For analysis of the understory model alone without tree shading (full sun), the light interception term $f_{St}(n)$ is set to 0 for these simulations.

b. Determine a priori conditions on parameters. All parameters are positive real. Parameter $S_2 > S_1$, as given in model description.

c. The model is implemented on a computational platform. Here the model is implemented in MATLAB 2018a.

d. If model equations are not twice continuously differentiable with respect to the parameters of interest, acceptable differentiable analytic approximations are available. There are two discontinuous functions in the model description. These are approximated with continuous equations as shown below.

e. General class of model inputs is known and ample input data at the model time step can be accessed. Inputs are synthesized by climate models and accessed through CliPick (see Section 3.1.3).

f. Acceptable a priori parameter values and initial conditions are available. The Burgess et al. (2014) parameterizations are used as the actual parameters

for generation reference outputs and ±5% variations of those parameters for *a priori* test parameterizations.

g. Cost function is defined. The standard sum squared error is used.

h. Measure of fit function is defined, and a cut-off criterion for goodness-of-fit selected. NRMSE is used with a cut-off threshold of 99% for an acceptable fit.

i. A measure of consistent identifiability for parameters is defined. We define consistent identifiability as optimized parameters from different *a priori* locations converge to within 0.01% of the same location.

i. A parameter optimization routine is implemented. The Trust-Region Reflective Newton algorithm as implemented in MATLAB via the `lsqnonlin` function was used (see Section 4.1.1).

Step 1. Compare goodness-of-fit of y_0 with the estimated output y_1.

Input u_{11} is used with actual parameter set annual grass to generate y_0, and *a priori* parameter set $\boldsymbol{\theta}_1$ generating the output realization y_1 (see Table 3.24 for parameter values). There was one adjustment made a parameter to facilitate the analysis. T_0 has an actual value of 0, which does not lend itself to the percentage difference analysis. Therefore, its actual value was set to 0.5°C for this analysis. As seen in Figure 3.19, the *a priori* model fit is unacceptable, with a NRMSE of 70.41%.

Table 3.24 Actual parameters θ_0 *a priori* settings θ_1 for crop model with other parameters set to those in Table 3.4.

Parameter	θ_0	θ_1
ε_c	0.3	0.285
T_0	0.5	0.475
P_0	0.8	0.76
S_1	1600	1520
S_2	1840	1748
k_c	0.4	0.38

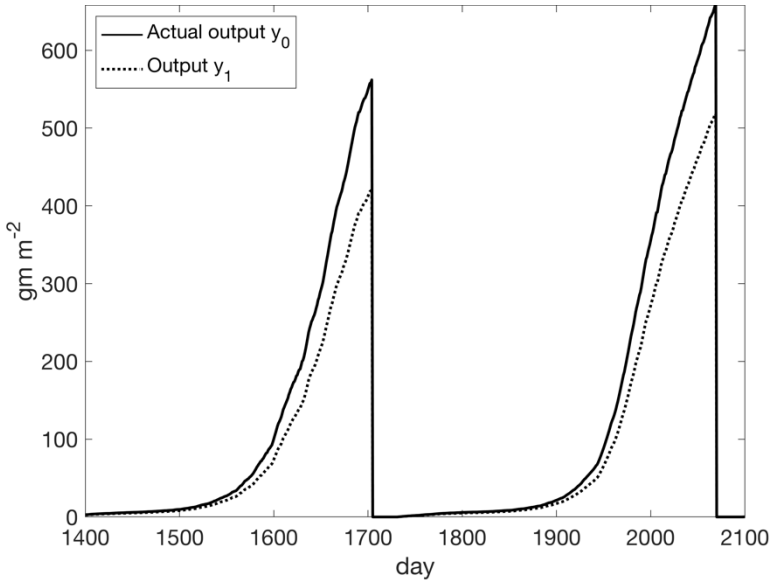

Figure 3.19 Outputs y_0 and y_1 for crop model driven by $u_{1,1}$ (n = 1400–2100). The NRMSE is 70.41%, indicating parameter must be adjusted to achieve an acceptable fit.

Step 2. Compute $H_{i,j}$ and its eigenvectors and eigenvalues.

The understory model was described in Section 3.1.1.2, and as noted previously, there are two discontinuous functions included in the descriptive equations. In order to carry out Hessian analysis, the system model needs to be twice continuously differentiable with respect to the parameters of interest. Therefore,

the discontinuous expressions for $P(n)$ and $S(n)$ are replaced with continuous, differentiable functions based on a commonly used approximation for the max function

$$\max(x_1, x_2, ..., x_m) \approx \frac{1}{\gamma} \ln(\sum_{i=1}^{m} e^{\gamma x_i}) \tag{3.18}$$

where γ is a large constant (but not too large in order to avoid numerical problems due to the $e^{\gamma x_i}$ term). $P(n)$ in (3.9) can be expressed as

$$\begin{aligned} P(n) &= P_0 - \min\left(P_0, \max\left(0, \frac{P_0(S(n) - S_1)}{S_2 - S_1} \right) \right) \\ &= P_0 + \max\left(-P_0, -\max\left(0, \frac{P_0(S(n) - S_1)}{S_2 - S_1} \right) \right) \end{aligned} \tag{3.19}$$

and since P_0 is a positive constant this can be rewritten as

$$P(n) = P_0 \left(1 + \max\left(-1, -\max\left(0, \frac{S(n) - S_1}{S_2 - S_1} \right) \right) \right) \tag{3.20}$$

Using (3.18) yields a differentiable approximation for $P(n)$

$$\begin{aligned} \tilde{P}(n) &= P_0 \left(1 + \frac{1}{\gamma} \ln\left(e^{-\gamma} + e^{-\gamma \frac{1}{\gamma} \ln\left(e^0 + e^{\frac{\gamma(S(n)-S_1)}{S_2-S_1}} \right)} \right) \right) \\ &= P_0 \left(1 + \frac{1}{\gamma} \ln\left(e^{-\gamma} + \left(1 + e^{\frac{\gamma(S(n)-S_1)}{S_2-S_1}} \right)^{-1} \right) \right) \end{aligned} \tag{3.21}$$

A differentiable approximation for $S(n)$ in (3.10) is also derived using (3.18)

$$\begin{aligned} \tilde{S}(n) &= \tilde{S}(n-1) + \frac{1}{\gamma} \ln(e^0 + e^{\gamma(T(n) - T_0)}) \\ &= \tilde{S}(n-1) + \frac{1}{\gamma} \ln(1 + e^{\gamma(T(n) - T_0)}) \end{aligned} \tag{3.22}$$

Comparison plots of the approximations in (3.21) and (3.22) are shown in Figure 3.20 and Figure 3.21, respectively.

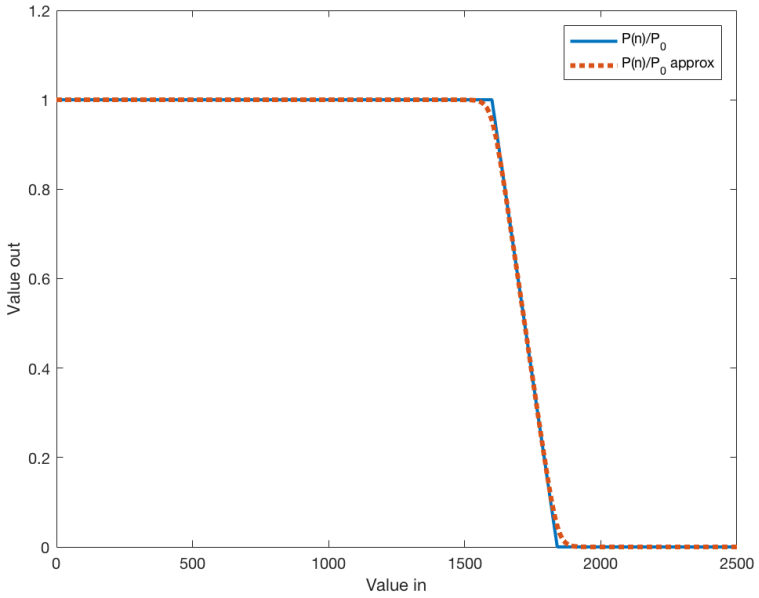

Figure 3.20: Comparison of $P(n)$ (3.20) to the approximation $\tilde{P}(n)$ (3.21) (divided by P_0) with γ=15.

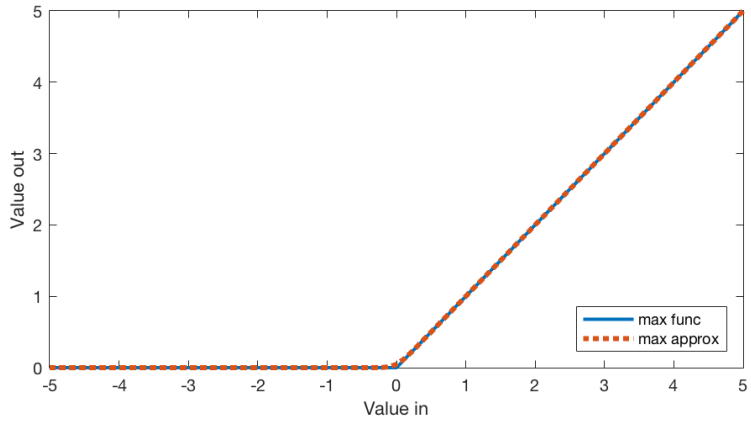

Figure 3.21: Comparison of the max function and its approximation used in $\tilde{S}(n)$ (3.22) with γ=15.

For convenience of notation, from here forward $P(n) \equiv \tilde{P}(n)$ and $S(n) \equiv \tilde{S}(n)$. Substituting (3.21) into (3.7) we arrive at the system under consideration for analysis

$$B_c(n) = B_c(n-1) + (1 - f_t(n))I(n)\varepsilon_c\left(1 - e^{-k_c L_c(n-1)}\right) \tag{3.23}$$

$$L_c(n) = L_c(n-1) - \Delta B_c(n) P_0 \sigma \left(1 + \frac{1}{\gamma}\ln\left(e^{-\gamma} + \left(1 + e^{\frac{\gamma(S(n)-S_1)}{S_2-S_1}}\right)^{-1}\right)\right) \tag{3.24}$$

$$S(n) = S(n-1) + \frac{1}{\gamma}\ln(1 + e^{\gamma(T(n)-T_0)}) \tag{3.25}$$

The revised form (3.24) makes it clear that P_0 and σ cannot both be estimated, as noted previously. This gives a fully defined model depicted in Figure 3.22 that is continuously differentiable in the 6 model parameters of interest ε_c, T_0, P_0, S_1, S_2, and k_c, with the variable γ set to a predetermined constant.

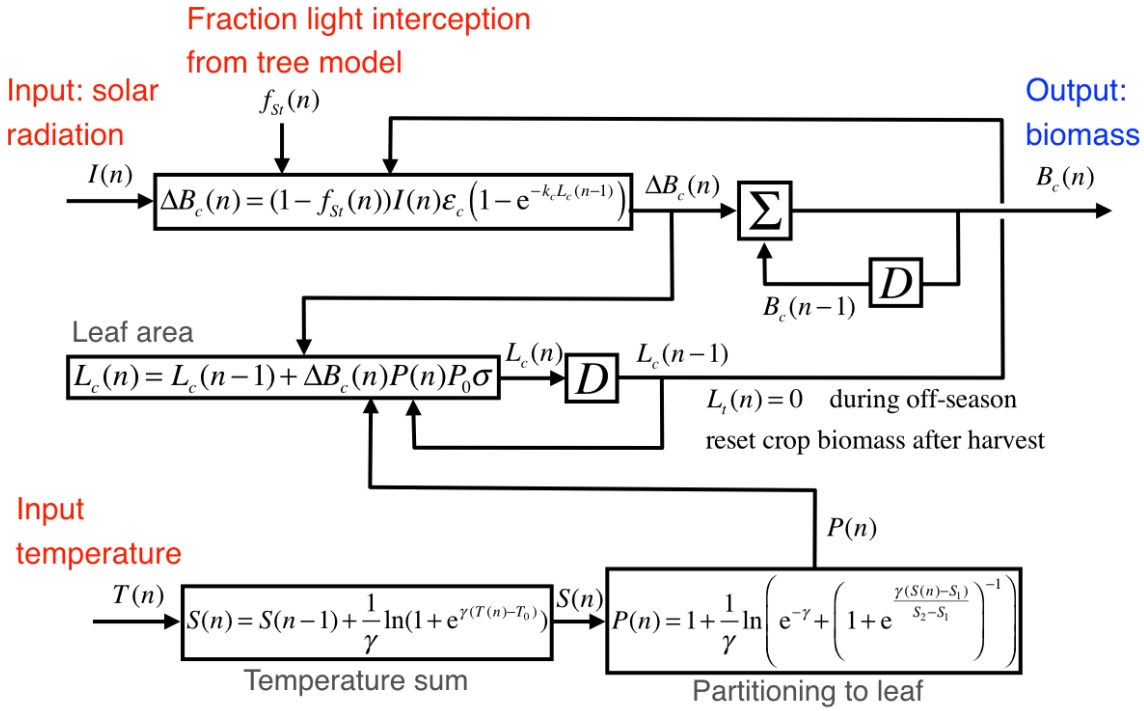

Figure 3.22: Approximate continuously differentiable revision of crop model. A user selected variable γ has been added for the min/max function approximations.

Combining equations (3.23), (3.24), and (3.25), into one yields a twice differentiable expression

$$B_c(i) = B_c(i-1) + (1-f_t(i))I(i)\varepsilon_c \times \left(1 - e^{-k_c\left(L_c(i-2) - \Delta B_c(i-1)P_0\sigma\left(1+\frac{1}{\gamma}\ln\left(e^{-\gamma}+\left(1+e^{\frac{\gamma\left(\left(S(i-2)+\frac{1}{\gamma}\ln(1+e^{\gamma(T(i-1)-T_0)})\right)-S_1\right)}{S_2-S_1}}\right)^{-1}\right)\right)\right)}\right) \quad (3.26)$$

Following the methods used for the tree model, parameter scaling prior to differentiation employed symbolic sidecar variables (Table 3.25) and the analytic Hessian with respect to the six parameters of interest was derived using MATLAB's `hessian` function. The Hessian calculated for N=7306 with eigenvalues and eigenvectors is given in Table 3.26. Three of the eigenvalues of the Hessian are grouped together within three orders of magnitude, while the smallest is insignificantly small (6.57e-12). Note that once again the Hessian has negative eigenvalues, indicating that the cost function surface is nonconvex.

Table 3.25 Just as for the tree model, crop model parameters are paired with symbolic variables prior to differentiation for Hessian. Parameter scaling is achieved by setting the model parameters to their *a priori* values and differentiating with respect to the symbolic variables s_1, s_2, ... s_6, then setting the symbolic variable values to 1.

Parameter		Parameter with symbolic variable
ε_c	→	$s_1\varepsilon_c$
T_0	→	s_2T_0
P_0	→	s_3P_0
S_1	→	s_4S_1
S_2	→	s_5S_2
k_c	→	s_6k_c

Table 3.26 Hessian and eigenvalues and eigenvectors for θ_1 and input $u_{1,1}$, ordered by eigenvalue magnitude.

	Hessian (N=7306)					
	9.72e-01	-7.63e-05	-4.52e-02	-9.01e-02	-1.77e-01	-2.12e+01
	-7.63e-05	-9.92e-08	-8.14e-05	-1.39e-04	-2.05e-04	-3.64e-05
	-4.52e-02	-8.14e-05	4.58e-04	-9.66e-02	-1.88e-01	-2.571e-02
	-9.01e-02	-1.39e-04	-9.66e-02	-3.39e-01	-1.22e-01	-4.50e-02
	-1.77e-01	-2.05e-04	-1.88e-01	-1.22e-01	-6.16e-01	-8.23e-02
	-2.12e+01	-3.64e-05	-2.57e-02	-4.50e-02	-8.23e-02	1.12e+01
	Eigenvalue					
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
	2.79e+01	-1.57e+01	-7.20e-01	-2.94e-01	6.31e-02	6.57e-12
	Eigenvector					
Axis	X_1	X_2	X_3	X_4	X_5	X_6
k_c	0.7857	0.6186	0.0104	0.0009	-0.0006	0.0000
ε_c	-0.6187	0.7856	0.0104	0.0010	-0.0004	0.0000
S_2	0.0016	0.0127	-0.8932	0.3815	0.2376	-0.0003
S_1	0.0007	0.0065	-0.3517	-0.9225	0.1587	-0.0003
P_0	0.0003	0.0035	-0.2797	-0.0582	-0.9583	-0.0000
T_0	0.0000	0.0000	-0.0004	-0.0002	0.0001	1.0000

Step 3. Generate Λ parameter ranking.

Based on Table 3.26, the Λ values and associated parameter ranking for this *a priori* setting are shown in Table 3.27.

Table 3.27 Λ values and their associated ranking based upon Hessian eigenstructure..

Ranking	**Parameter**	Λ
1	k_c	3.1612e+01
2	ε_c	2.9582e+01
3	S_2	1.0134e+00
4	S_1	6.5738e-01
5	P_0	3.4095e-01
6	T_0	4.1980e-04

This leads to the following $\chi_{1,k}$ matrix (Table 3.28) of candidate parameter vectors for estimation.

Table 3.28 $\chi_{1,k}$ candidates for reduced-order parameter estimation crop model.

	$\chi_{1,1}$	$\chi_{1,2}$	$\chi_{1,3}$	$\chi_{1,4}$	$\chi_{1,5}$	$\chi_{1,6}$
k_c	X	X	X	X	X	X
ε_c		X	X	X	X	X
S_2			X	X	X	X
S_1				X	X	X
P_0					X	X
T_0						X

Step 4. Confirm consistent identifiability.

$\chi_{1,1}$ indicates that k_c is the first parameter to estimate. Optimizations from both *a priori* settings at $\pm 5\%$ of the actual k_c value (θ_1 and θ_{33}) converge to the same estimate, as shown in Table 3.29.

Table 3.29 Actual parameters and $\pm 5\%$ *a priori* settings θ_1 and θ_2 for the first ranked candidate parameter R together with the optimized setting. For comparison, θ_0 is also included.

Parameter	θ_1	θ_{33}	$\hat{\theta}_1$ **and** $\hat{\theta}_{33}$	θ_0
k_c	0.38	0.42	0.4539	0.4
ε_c	0.285	0.285	0.285	0.3
S_2	1748	1748	1748	1840
S_1	1520	1520	1520	1600
P_0	0.76	0.76	0.76	0.8
T_0	0.475	0.475	0.475	0.5

Step 5. Check goodness-of-fit of y_0 and $\hat{y}_j$.

The NRMSE value for the model with reduced parameter space $\chi_{1,1}$ estimated is 96.26%, which does not meet the criterion of ≥99%. Return to Step 4 with the next candidate parameter vector.

Step 4 (Attempts 2–3) Confirm consistent identifiability.

For brevity, Step 4 attempts 2–3 are condensed into Table 3.30. $\chi_{1,3}$ is shown to meet the fit criteria and be identifiable from all combinations of ±5% of each of its estimated parameters.

Table 3.30 Step 4 for the first three candidate parameter sets.

	$\chi_{1,1}$	$\chi_{1,2}$	$\chi_{1,3}$	Actual
	$\hat{\theta}_1$, $\hat{\theta}_{33}$	$\hat{\theta}_1$, $\hat{\theta}_2$, $\hat{\theta}_{33}$, $\hat{\theta}_{34}$	$\hat{\theta}_1$, $\hat{\theta}_2$, $\hat{\theta}_{33}$, $\hat{\theta}_{34}$, $\hat{\theta}_{17}$, $\hat{\theta}_{18}$, $\hat{\theta}_{49}$, $\hat{\theta}_{50}$	θ_0
Max % param. diff.	7.0752e-07%	1.7387e-03%	2.9926e-04%	–
k_c	0.4539	0.46307	0.41116	0.4
ε_c	0.285	0.28036	0.30352	0.3
S_2	1748	1748	1945	1840
S_1	1520	1520	1520	1600
P_0	0.76	0.76	0.76	0.8
T_0	0.475	0.475	0.475	0.5
NRMSE	96.26%	96.27%	99.15%	100%

Step 6. Validation.

As shown in Table 3.31, the model fit for $\hat{\theta}_1$ in $\chi_{1,3}$ remains above the NRMSE threshold for all validation sequences.

Table 3.31 Fit measures for $\hat{\hat{\theta}}_1$ for $\chi_{1,3}$.

Input	Use	NRMSE	Cost gm^2 m^{-2}
$u_{1,1}$	Optimization	99.15%	0.63
$u_{1,2}$	Validation	99.42%	0.66
$u_{1,3}$	Validation	99.30%	1.04
$u_{1,4}$	Validation	99.22%	0.60
$u_{1,5}$	Validation	99.15%	0.62

Step 4 (Optional step: check $\chi_{1,4}$) Confirm consistent identifiability.

$\chi_{1,3}$ led to an acceptable model fit. In this optional additional test, we investigate if the fit improves for $\chi_{1,4}$. Table 3.32 shows that $\chi_{1,4}$ is consistently identifiable.

Table 3.32 Step 4 for $\chi_{1,4}$.

	$\chi_{1,4}$	Actual
	$\hat{\theta}_1$, $\hat{\theta}_j$ (j=2, 9, 10, 17, 18, 25, 26, 33, 34, 41, 42, 49, 50, 57, 58)	θ_0
Max % param. diff.	1.4905e-04%	–
k_c	0.40207	0.4
ε_c	0.30758	0.3
S_2	1851	1840
S_1	1609	1600
P_0	0.76	0.8
T_0	0.475	0.5
$NRMSE$	99.73%	100%

Step 5. (4 parameters) Check goodness-of-fit of y_0 and $\hat{y}_j$.

As shown in Table 3.33, the model fit improves with estimation of 4 parameters.

Table 3.33 Fit measures for $\hat{\hat{\theta}}_1$ for $\chi_{1,4}$ for L_1 validation inputs.

Input	Use	NRMSE	Cost gm^2 m^{-2}
$u_{1,1}$	Optimization	99.73%	0.64
$u_{1,2}$	Validation	99.60%	0.32
$u_{1,3}$	Validation	99.40%	0.75
$u_{1,4}$	Validation	99.73%	0.07
$u_{1,5}$	Validation	99.81%	0.03

3.3.1.1 Simulations to show robustness of HIROPE for the crop model

First, performance for the accepted $\hat{\hat{\theta}}_1$ parameterization is evaluated for inputs from other locations. The model was run with inputs from locations L_2–L_{10} to compare outputs for parameter settings at θ_0 and $\hat{\hat{\theta}}_1$. Fit measures are given in Table 3.34. Unlike the tree model, performance falls below the 99% threshold for several locations. However, all NRMSE values are >98.42%, which is a respectable fit for locations other than that for which the fit was optimized.

Table 3.34 Fit measures for $\hat{\hat{\theta}}_1$ for $\chi_{1,4}$ using inputs from locations other than the location used for parameter estimation.

		$\hat{\hat{\theta}}_1$ for $\chi_{1,4}$	
Input	Use	NRMSE	Cost gm^2 m^{-2}
$u_{1,1}$	Optimization	99.73%	0.64
$u_{2,1}$	Validation	99.46%	0.43
$u_{3,1}$	Validation	98.87%	14.30
$u_{4,1}$	Validation	99.37%	0.99
$u_{5,1}$	Validation	98.42%	26.14
$u_{6,1}$	Validation	99.04%	3.08
$u_{7,1}$	Validation	99.40%	0.73
$u_{8,1}$	Validation	98.87%	4.74
$u_{9,1}$	Validation	98.75%	10.04
$u_{10,1}$	Validation	98.78%	7.70

3.3.1.2 *Ranking by* Λ *for all 64* θ_j

As a test of robustness of the Λ ranking, the Λ_i are calculated for all parameters and all 64 θ_j. As shown in Table 3.35, the order of importance of the parameters is consistent, although not quite as consistent as for the tree model. The relative ranking of the top two parameters is the same at all θ_j. Figure 3.23 and Figure 3.24 show Λ_i for all 64 θ_j.

Table 3.35 Number of occurrences of parameter ranking for all 64 *a priori* estimates θ_j.

	Ranking					
Par.	**1**	**2**	**3**	**4**	**5**	**6**
k_c	59	5	0	0	0	0
ε_c	5	59	0	0	0	0
S_2	0	0	43	19	2	0
S_1	0	0	21	43	0	0
P_0	0	0	0	2	62	0
T_0	0	0	0	0	0	64

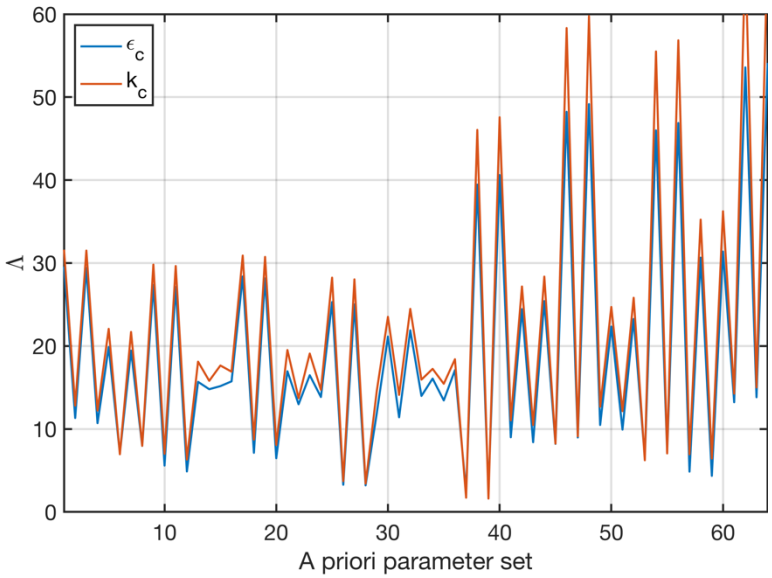

Figure 3.23 Λ_i for the two highest ranking parameters and all 64 θ_j.

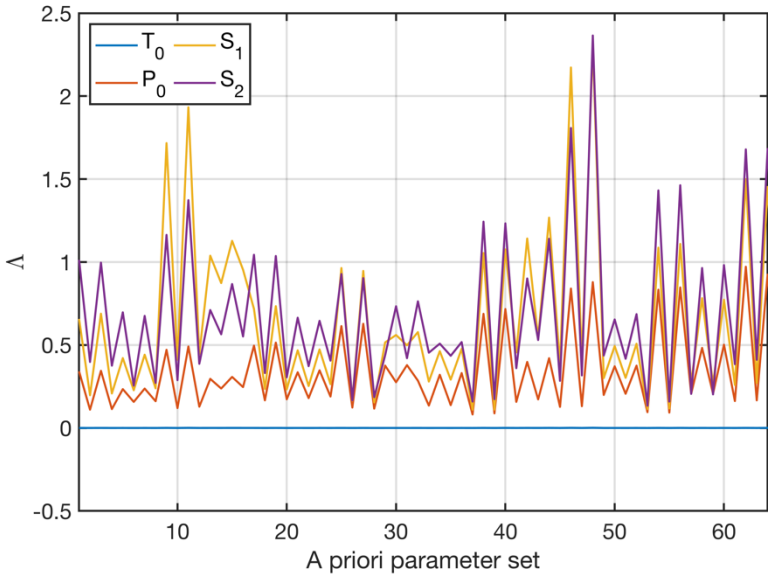

Figure 3.24 Λ_i for the four lowest ranking parameters and all 64 θ_j. Note: these Λ values are an order of magnitude smaller than for the two highest ranked parameters.

3.4 Tree and crop model combination (non-water limiting case)

Reduction of the Yield-SAFE tree and crop combination model follows the analysis conducted for each above. The primary difference is that the understory model description requires substitution of two model equations with continuously differentiable approximations, as detailed below.

For the following analysis of the tree model, the following are used

Parameter vector $\Theta = [\ \varepsilon_t\ \alpha\ L_m\ B_t(0)\ N(0)\ R\ \varepsilon_c\ T_0\ P_0\ S_1\ S_2\ k_c]$

Actual parameters:

$$\theta_0 = [\text{Poplar1 (from Table 3.2) Annual grass (from Table 3.4)}]$$

There are 2^{12} or 4096 $\pm 5\%$ variations in *a priori* parameters θ_j (ordered following the pattern established in Table 3.7).

3.4.1 Advantages of a reduced-order parameter space for estimation

As a reminder, one would wish to have a parameter estimation procedure that reliably locates a parameter estimates with acceptable model performance from any reasonable initial parameter estimate. Before demonstrating the HIROPE procedure on the 12-parameter combination model and to show why locating a reduced-order parameter space is important for complex, nonlinear growth models such as Yield-SAFE, it is revealing to run an experiment where all parameters are estimated simultaneously. Section 3.1.4.2 demonstrated that for the tree model alone, attempts to estimate six of its parameters simultaneously revealed a lack of consistent identifiability. In that case, there existed a manifold in R–N(0) space where the cost function was nearly zero as long as the other four parameters were very close to their actual values. In finding a reduced parameter space for estimation, the HIROPE procedure imposes a requirement that the optimization problem have a unique solution in the parameter space (defined as consistent identifiability in Section 2.4).

If one were to relax the requirement for consistent identifiability for the combination tree-crop model, what model performance can be achieved by estimation of all parameters in the perfect identification case (identical model and identifier structure and noise-free measurements, Figure 3.12)? For the answer to this question, the 12-parameter estimation was conducted from 100 *a priori* parameter settings selected at random using MATLAB's `randperm` function from the 4096 ±5% variations of $\boldsymbol{\theta}_0$. The input is again $u_{1,1}$ (Norwich, UK). Table 3.36 lists the index numbers j for the randomly selected $\boldsymbol{\theta}_j$ used in these simulations, while Figure 3.25 depicts a histogram of the NRMSE values for the y_j.

For comparison, estimation of 4 of the 12 parameters was carried out for the same θ_j. The parameters selected for the 4-parameter estimation example were those found to give acceptable model performance when estimated for the individual tree and crop models in previous sections: ε_t, ε_c, S_2, and k_c. The eight other parameters were fixed at their *a priori* θ_j values.

Table 3.36 Indices (j) for the 100 θ_j used for this simulation. The ±5% variations in the θ_j (and therefore their index numbers) follow the pattern established for the 6-parameter models in Table 3.7.

2	61	71	86	93	124	144	152	195	209
211	288	364	375	380	429	454	467	516	597
598	614	663	686	798	800	814	845	851	873
903	1010	1049	1083	1133	1188	1203	1257	1291	1308
1344	1428	1464	1527	1567	1586	1644	1657	1761	1790
1841	1852	1911	2044	2069	2127	2165	2169	2229	2360
2393	2466	2506	2619	2621	2693	2720	2727	2772	2883
2924	3054	3076	3130	3138	3139	3140	3155	3173	3198
3341	3353	3402	3412	3437	3460	3523	3524	3542	3552
3638	3643	3688	3718	3725	3752	3839	3975	4025	4035

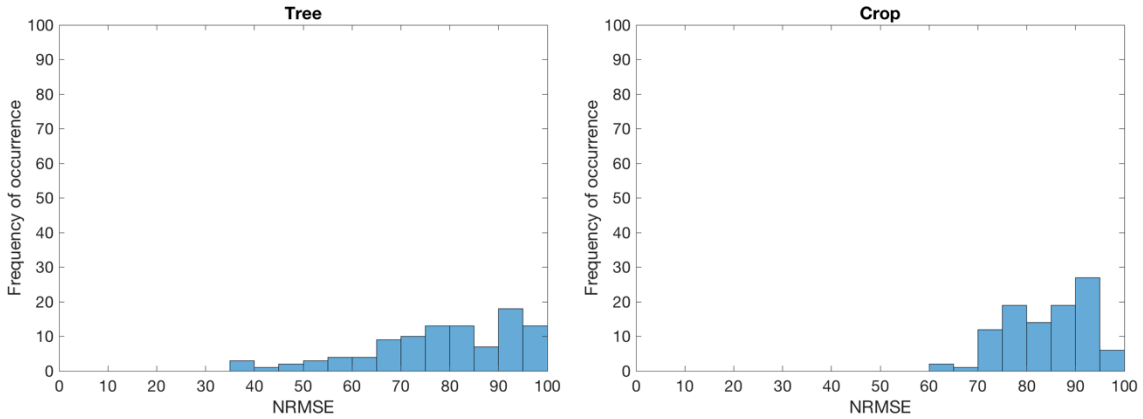

Figure 3.25 Histograms of the NRMSE values for the tree biomass (left) and the crop biomass (right) for the 100 θ_j used for this simulation.

The same nonlinear least squares algorithm used for the individual tree and crop models (the `lsqnonlin` option in `nlgreyest`) was utilized for the combination

model, as well as the same termination criterion on the lower bound of change in the cost function (`TolFun` = 10^{-5}). The maximum number of iterations of the algorithm was set to 150, after which algorithm progress was considered to have stalled. A high number of iterations suggests that the parameter estimate is caught in a region of the parameter space where changes in parameters have only a very small effect on the cost function.

Table 3.37 lists the frequency of the termination causes from the 100 *a priori* settings. For the 12-parameter estimation, 69% of estimations terminated when the change in the cost function for the next iteration fell below the lower bound of 10^{-5}. In other words, the algorithm was unable to locate a new parameter estimate in the vicinity of the last estimate that decreased the cost function more than 10^{-5}. In 31% of estimations, termination was due to exceeding 150 iterations. In contrast to the 12-parameter case, nearly all estimations terminated due to the cost function change bound in the 4-parameter estimation case.

Table 3.37 Frequency of termination causes of the nonlinear least squares algorithm from 100 θ_j used for this simulation.

	Termination condition	
	Lower bound of change in the cost function `TolFun`$<10^{-5}$	**# iterations >150**
12-parameter estimation	69	31
4-parameter estimation	99	1

For the 12-parameter estimation, histograms of NRMSE values for the optimized parameters in Figure 3.26 show that while for most $\hat{\theta}_j$ performance is greatly improved over that of the θ_j, 28% of estimated parameter settings result in unsatisfactory performance of the crop model (NRMSE$<$95). As can be seen from

the parameter estimates of five example $\hat{\theta}_j$ in Table 3.38, the values varied greatly depending upon their initializations. In about 20% of initializations, $\hat{\theta}_j$ was very close to θ_0 with exception of parameters R and N(0), which reached the afforementioned R–N(0) manifold, giving a nearly perfect fit to the reference outputs.

Table 3.38 Five example $\hat{\theta}_j$ and NRMSE values from the estimations carried out on all 12 parameters of the combination model.

	j index for $\hat{\theta}_j$					
Param	873	1083	1133	2165	2229	**θ_0**
ε_t	1.0351e+00	1.4110e+00	1.1341e+00	1.4086e+00	1.3661e+00	1.4086e+00
α	2.7849e-05	1.0020e-04	5.6831e-05	1.0000e-04	9.4880e-05	1.0000e-04
L_m	1.1453e+03	4.9762e+02	8.6683e+02	5.0000e+02	5.3865e+02	5.0000e+02
$B_t(0)$	9.9597e+01	1.0001e+02	1.0051e+02	9.5000e+01	9.3761e+01	1.0000e+02
N(0)	5.7059e-01	6.3828e-01	6.0728e-01	6.5373e-01	6.0144e-01	6.2250e-01
R	2.1000e+05	2.1001e+05	2.1000e+05	2.1131e+05	2.1096e+05	2.0000e+05
ε_c	3.0689e-01	3.0014e-01	1.1758e-01	2.9999e-01	2.9589e-01	3.0000e-01
T_0	2.4061e+00	1.9067e-01	6.1408e-01	4.3117e-01	1.5032e+00	5.0000e-01
P_0	8.9661e-01	7.9434e-01	8.1184e-01	8.0004e-01	7.9005e-01	8.0000e-01
S_1	1.9584e+03	1.7562e+03	2.0655e+03	1.6183e+03	1.6244e+03	1.6000e+03
S_2	1.9584e+03	1.8509e+03	2.0768e+03	1.8578e+03	7.0558e+03	1.8400e+03
k_c	4.2178e-01	4.0048e-01	8.3619e-01	4.0000e-01	4.4270e-01	4.0000e-01
Tree NRMSE	98.80	100	99.64	100	99.91	100
Crop NRMSE	82.33	99.65	33.19	100	96.67	100

One might consider the perfect fit outcomes in the 12-parameter case to be an advantage over reduced-order estimation, which generally cannot attain a perfect fit since some parameters are fixed to unfavorable values. However, in practice a model never matches the real system exactly and measurement data contains noise. Therefore, the possiblity that the 12-parameter estimation can reach the actual parameters and a perfect fit from some initializations is an artificial

consequence of the perfect identification set-up used here. Rather than the impractical goal of estimating the actual parameters, it is more important that the parameter estimation is reliable from any reasonable initialization. As shown in Figure 3.26, a substantial number of initalizations result in poor performance in the 12-parameter estimation case.

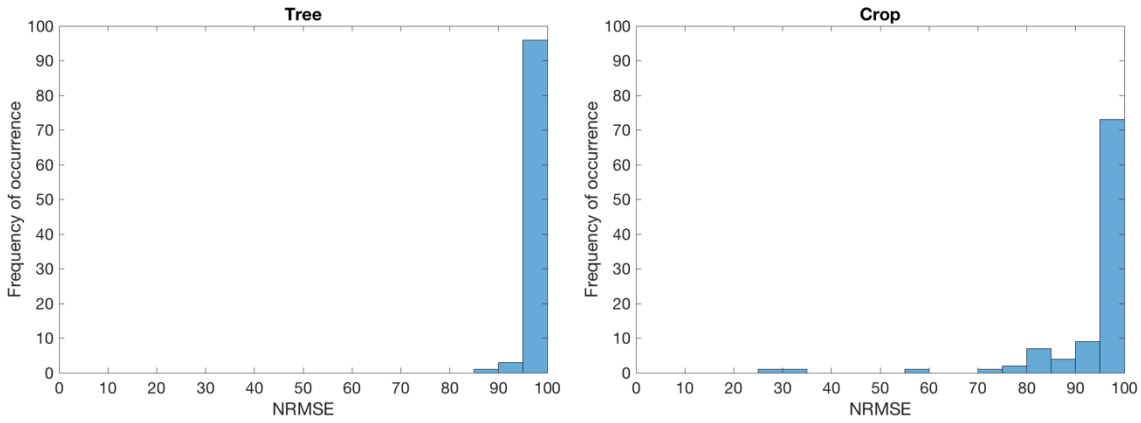

Figure 3.26 NRMSE values for the combination model parameterized by the 100 $\hat{\theta}_j$ generated by estimating all 12 parameters. The average NRMSE values and standard deviations are is 99.42±1.74 for the tree model and 93.41±15.29 for the crop model.

As shown in the 4-parameter example (Table 3.39 and Figure 3.27), the reduced-order estimation has a clear advantage in consistent performance from nearly all $\hat{\theta}_j$. In this case for the tree output the NRMSE values are >95 in all cases, and the average NRMSE is about the same as for the 12-parameter case (98.39±1.16). However, for 4-parameter estimation, the crop model performs significantly better on average than the 12-parameter case (97.18±4.71 vs. 93.41±15.29, respectively). In the 4-parameter case, only 6% have NRMSE values <95, as compared with 28% when estimating 12 parameters.

Table 3.39 Five example $\hat{\theta}_j$ and NRMSE values from the estimations carried out on 4 parameters (ε_t, ε_c, S_2, and k_c) of the combination model.

	j index for $\hat{\theta}_j$					
Param	873	1083	1133	2165	2229	**θ_0**
ε_t	1.5387e+00	1.4619e+00	1.4789e+00	1.3190e+00	1.3190e+00	1.4086e+00
α	9.5000e-05	1.0500e-04	9.5000e-05	9.5000e-05	9.5000e-05	1.0000e-04
L_m	4.7500e+02	4.7500e+02	5.2500e+02	5.2500e+02	5.2500e+02	5.0000e+02
$B_t(0)$	1.0500e+02	1.0500e+02	1.0500e+02	9.5000e+01	9.5000e+01	1.0000e+02
N(0)	5.9137e-01	6.5363e-01	5.9137e-01	6.5363e-01	6.5363e-01	6.2250e-01
R	2.1000e+05	2.1000e+05	2.1000e+05	2.1000e+05	2.1000e+05	2.0000e+05
ε_c	3.0707e-01	3.0943e-01	3.1285e-01	2.9217e-01	2.9211e-01	3.0000e-01
T_0	4.7500e-01	4.7500e-01	4.7500e-01	4.7500e-01	5.2500e-01	5.0000e-01
P_0	8.4000e-01	7.6000e-01	7.6000e-01	7.6000e-01	7.6000e-01	8.0000e-01
S_1	1.6800e+03	1.5200e+03	1.5200e+03	1.5200e+03	1.5200e+03	1.6000e+03
S_2	8.2365e-04	1.8705e+03	1.6591e+03	2.1981e+03	2.2490e+03	1.8400e+03
k_c	6.0995e-01	3.9447e-01	3.8556e-01	4.4120e-01	4.4132e-01	4.0000e-01
Tree NRMSE	9.6194e+01	9.9598e+01	9.6807e+01	9.8230e+01	9.8230e+01	100
Crop NRMSE	8.6508e+01	9.8738e+01	9.7243e+01	9.7125e+01	9.7153e+01	100

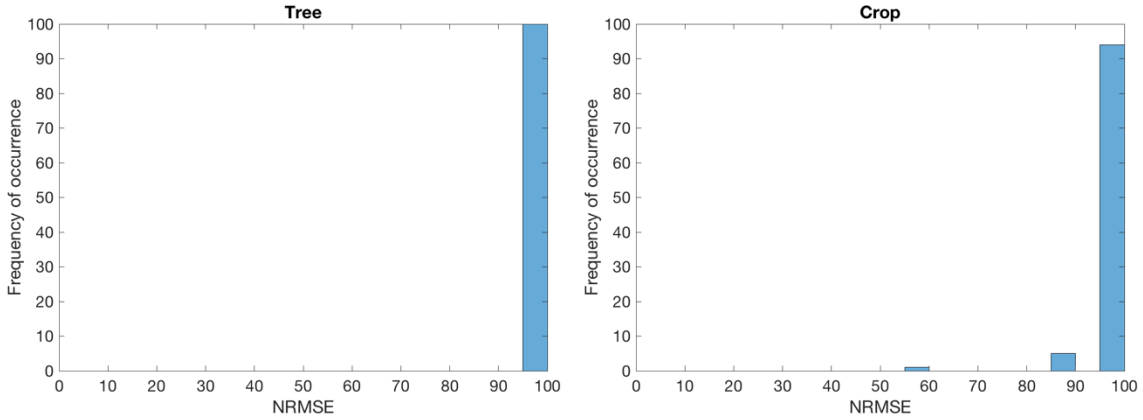

Figure 3.27 NRMSE values for the combination model parameterized by the 100 $\hat{\theta}_j$ generated from optimizing the 4 parameters (ε_t, ε_c, S_2, and k_c). The average NRMSE values and standard deviations are is 98.39±1.16 for the tree model and 97.18±4.71 for the crop model.

One may wonder if the relative performance of full- versus reduced-order parameter estimation changes if the initial parameter error is greater. As additional test, 12- and 4-parameter optimizations were run from initial settings ±10% away from the actual parameters (instead of ±5%). The results mirror the relative performance at ±5% error, as shown in Table 3.40, with the 4-parameter estimation performing well for both tree and crop outputs, while the 12-parameter estimation performs significantly worse for the crop output.

Table 3.40 Comparison of average NRMSE values for ±5% and ±10% in $\hat{\theta}_j$.

	Tree model output		**Crop model output**	
	12-par. estimation	**4-par. estimation**	**12-par. estimation**	**4-par. estimation**
$\hat{\theta}_j$ error	**NRMSE**	**NRMSE**	**NRMSE**	**NRMSE**
±5%	99.42±1.74	98.39±1.16	93.41±15.29	97.18±4.71
±10%	96.70±5.63	96.34±2.39	75.11±36.18	93.33±5.30

Finally, one might wonder how the 4-parameter estimation performs for another set of four parameters. To demonstrate the value of a procedure for ranking parameter importance to the cost function, the 4-parameter estimation was run on four parameters that were not selected based on the HIROPE procedure applied to the tree and crop models: α, S_1, S_2, and P_0. In this case, the estimated parameters lead to a improvement over the θ_0, although overall the resulting performance was much worse than achieved by estimating the four HIROPE-selected parameters. Histograms of the NRMSE values for this experiment are shown in Figure 3.28.

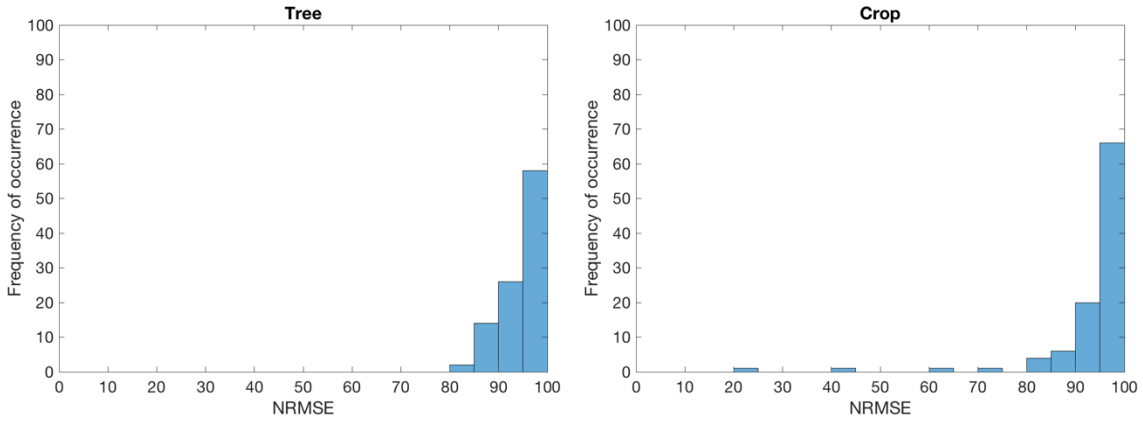

Figure 3.28 NRMSE values for the combination model parameterized by the 100 $\hat{\theta}_j$ generated from optimizing α, S_1, S_2, and P_0. The average NRMSE values and standard deviations are is 94.65±3.92 for the tree model and 93.35±10.34 for the crop model.

One might claim that a different optimization algorithm or cost function formulation might lead to better results. However, this simulation demonstrates the results that will likely arise for an end-user when attempting to estimate all model parameters versus a carefully selected parameter subset. These results also foreshadow even greater problems with estimating all model parameters as the dimension of the parameter space increases. In the following, the HIROPE procedure is applied to the full combination model.

3.4.2 HIROPE procedure applied to tree and crop model combination (non-water limiting case)

For this example, the HIROPE procedure is run using θ_1 (all parameters –5% away from the actual parameters). A list of procedure steps and expected outcomes are given in Table 3.41.

Table 3.41. Simulation list for combination model.

M_{tc} (combination tree and crop)	
Action	Expected outcome
Repeat analysis described above for M_t.	Because tree model affects the amount of light reaching the crop model, it is expected that the important tree model parameters will have higher ranking than the important crop model parameters, but that a good fit to both tree and crop model outputs will only be achieved by optimizing both tree and crop model parameters.
Before optimizing on combination model run combination model with parameter sets optimized individually for the overstory and understory models, and compute NRMSE values.	NRMSE values may be acceptable, but improvements may be gotten through optimization on the full model.

3.4.3 Prerequisites

The prerequisites for the combination model all follow the descriptions given for the tree and crop models, with two differences, as noted below

a. The full system equations are defined. Given in the tree and crop models, with the crop model receiving solar radiation reduced by $f_{St}(n)$, the light intercepted by the trees, as given by

$$f_{St}(n) = 1 - e^{-\rho k(n) L_t(n)} \tag{3.27}$$

b. Determine a priori conditions on parameters. Same as for tree and crop models.

c. The model is implemented on a computational platform. The model is implemented in MATLAB 2018a.

d. If model equations are not twice continuously differentiable with respect to the parameters of interest, acceptable differentiable analytic approximations are available. Same as for tree and crop models (including the substitution of differentiable functions in the crop model, Figure 3.29).

e. General class of model inputs is known and ample input data at the model time step can be accessed. Same as for tree and crop models.

f. Acceptable a priori parameter values and initial conditions are available. Same as for tree and crop models.

g. Cost function is defined. A standard sum squared error is used, with the difference that this model has two outputs and a scaling factor between the two outputs (s_{comb}). The cost function becomes

$$V_N(n) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}\left(r_t^2(i) + s_{comb}r_c^2(i)\right) \tag{3.28}$$

where the indices t and c refer to tree and crop models, respectively. Because the tree and crop submodel outputs are in different units and are scaled differently, their values are normalized with a scaling factor by calculating the ratios of the average error magnitude between the actual and *a priori* outputs. The scaling factor is

$$s_{comb} = \frac{\left\|y_{t0} - y_{t1}\right\|_1}{\left\|y_{c0} - y_{c1}\right\|_1} \tag{3.29}$$

For θ_0 and model inputs $u_{1,1}$, s_{comb} is 55.16, which multiplies the squared error of the crop model output for the Hessian calculation.

h. Measure of fit function is defined, and a cut-off criterion for goodness-of-fit selected. Same as for tree and crop models, NRMSE >99%.

j. A measure of consistent identifiability for parameters is defined. Same as for tree and crop models.

i. A parameter optimization routine is implemented. Same as for tree and crop models.

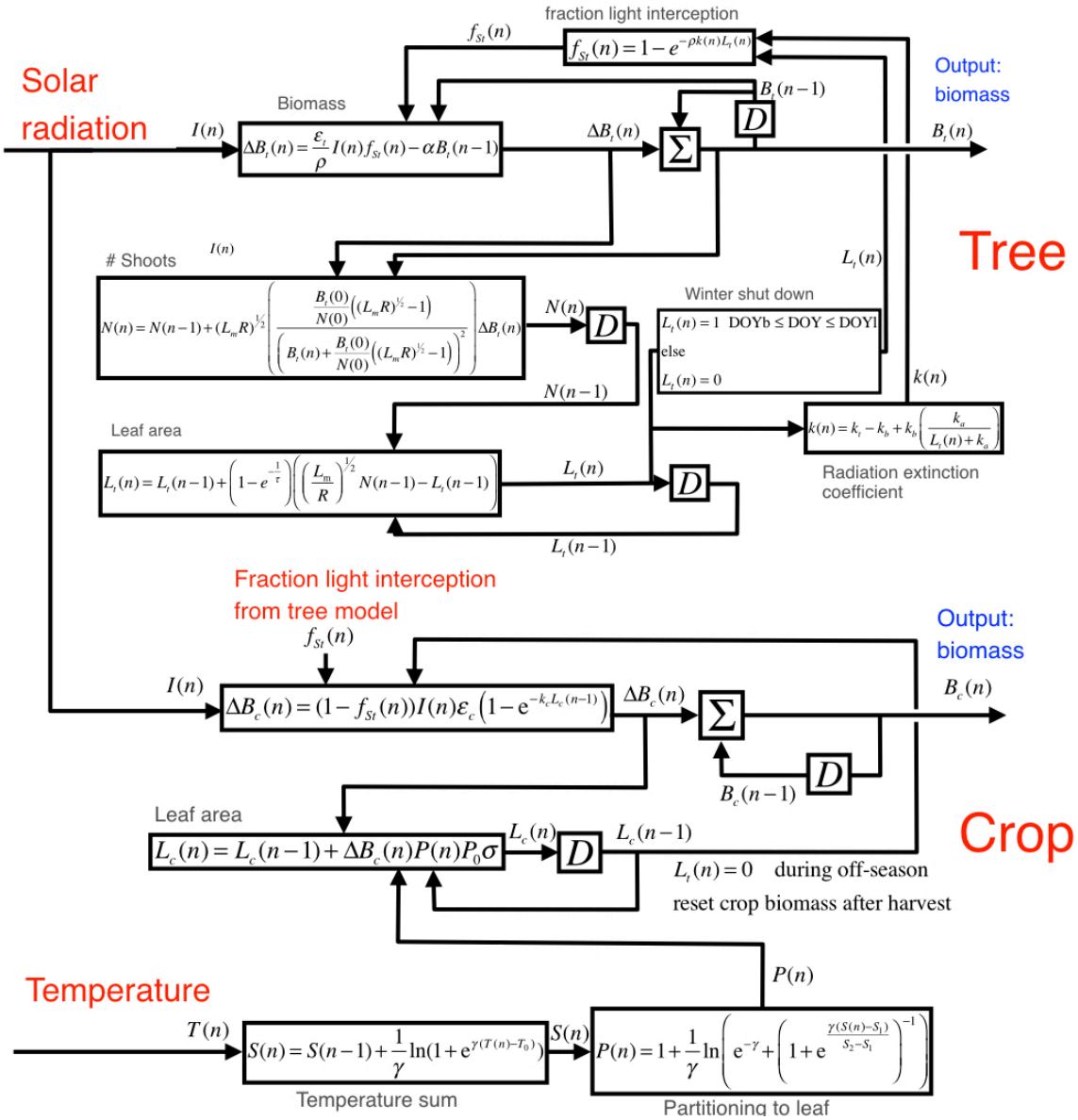

Figure 3.29 Combined tree and crop model with differentiable function replacements in crop model.

Step 1. Compare goodness-of-fit of y_0 with the estimated output y_1.

For this 2-output model, each output is compared with the reference (i.e., y_{t0} with y_{t1} and y_{c0} with y_{c1}, see parameter values in Table 3.42). The tree output is not affected by the crop output, while the crop output is affected by the tree output

through the connecting light interception term. Therefore, the fit of the tree model is 79.09% as before, while the crop model fit is 73.45% (Figure 3.30). In the case of θ_1, the fit of the understory is slightly better than for the crop model alone, however, the crop model fit will in general be worse for the combination model because error in the tree model can compound error in the crop model.

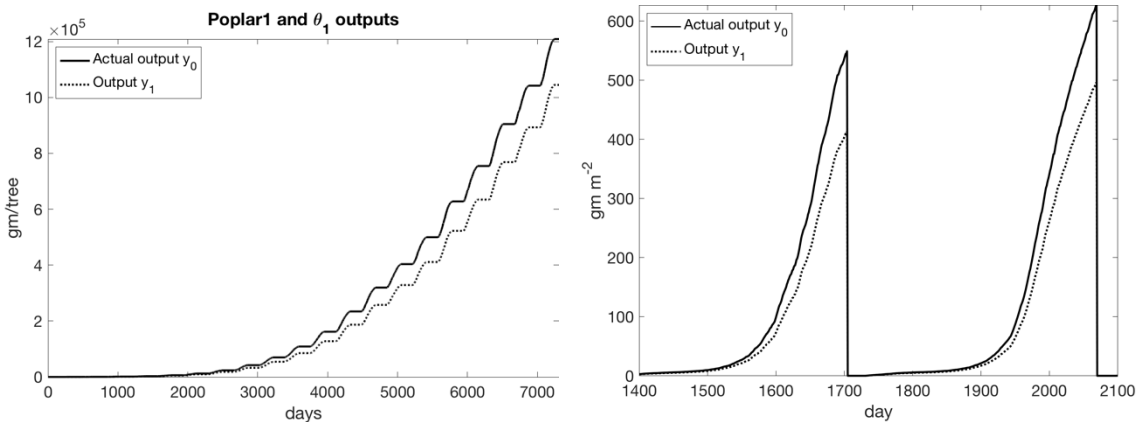

Figure 3.30 Outputs y_0 and y_1 for tree (left) and crop (right, n = 1400–2100) models driven by $u_{1,1}$. The NRMSE values are 79.09% and 73.45%, respectively, indicating parameter must be adjusted to achieve an acceptable fit.

Table 3.42 Actual parameters of interest θ_0 and *a priori* settings θ_1 for combination model.

Tree Parameters	θ_0	θ_1	Crop Parameters	θ_0	θ_1
ε_t	1.4086	1.33817	ε_c	0.3	0.285
α	0.0001	0.000095	T_0	0.5	0.475
L_m	500	475	P_0	0.8	0.76
$B_t(0)$	100	95	S_1	1600	1520
N(0)	0.6225	0.591375	S_2	1840	1748
R	200000	190000	k_c	0.4	0.38

Step 2. Compute $H_{i,j}$ and its eigenvectors and eigenvalues.

Following the methods used for the tree and crop model, parameter scaling prior to differentiation employed symbolic sidecar variables and the analytic Hessian with respect to the 12 parameters of interest was derived using MATLAB's

`hessian` function. The Hessian calculated for N=5475 (15 years) is shown in Table 3.43, with its eigenvalues and eigenvectors given in Table 3.44.

Table 3.43 Hessian for θ_1 and input $u_{1,1}$, combination system.

Hessian (N=5475)					
Column 1	Column 2	Column 3	Column 4	Column 5	Column 6
3.58e+05	-1.85e+04	-3.48e+05	5.16e+02	-5.16e+02	3.486e+05
-1.85e+04	1.25e+03	-3.09e+02	4.78e-01	-4.78e-01	3.09e+02
-3.48e+05	-3.09e+02	1.89e+05	4.66e+02	-4.66e+02	1.64e+05
5.16e+02	4.78e-01	4.66e+02	-6.11e+02	8.67e+01	-2.91e+01
-5.16e+02	-4.78e-01	-4.66e+02	8.67e+01	4.37e+02	2.91e+01
3.48e+05	3.09e+02	1.64e+05	-2.91e+01	2.91e+01	-5.17e+05
-1.56e-03	5.88e-05	7.80e+02	-3.01e+00	3.01e+00	-7.79e+02
-6.16e-09	2.05e-10	2.21e-03	-1.09e-05	1.09e-05	-2.21e-03
-6.81e-06	2.35e-07	2.88e+00	-1.30e-02	1.30e-02	-2.87e+00
-9.94e-06	3.34e-07	3.60e+00	-1.77e-02	1.77e-02	-3.60e+00
-1.12e-05	3.69e-07	3.99e+00	-1.98e-02	1.98e-02	-3.99e+00
-9.19e-04	3.61e-05	4.47e+02	-1.54e+00	1.54e+00	-4.47+02
Column 7	Column 8	Column 9	Column 10	Column 11	Column 12
-1.56e-03	-6.16e-09	-6.81e-06	-9.94e-06	-1.12e-05	-9.19e-04
5.88e-05	2.05e-10	2.35e-07	3.34e-07	3.69e-07	3.61e-05
7.80e+02	2.21e-03	2.88e+00	3.60e+00	3.99e+00	4.47e+02
-3.01e+00	-1.09e-05	-1.30e-02	-1.77e-02	-1.98e-02	-1.53e+00
3.01e+00	1.09e-05	1.30e-02	1.77e-02	1.98e-02	1.54e+00
-7.79e+02	-2.21e-03	-2.87e+00	-3.60e+00	-3.99e+00	-4.47e+02
9.18e+03	-4.35e-01	-6.90e+02	-7.42e+02	-7.49e+02	-7.73e+04
-4.35e-01	1.62e-04	-4.57e-01	2.19e-01	3.44e-01	2.19e-01
-6.90e+02	-4.57e-01	1.63e+01	-7.77e+02	-7.90e+02	-8.15e-01
-7.42e+02	2.19e-01	-7.77e+02	1.44e+02	6.39e+02	3.08e+02
-7.49e+02	3.44e-01	-7.90e+02	6.39e+02	5.30e+02	4.50e+02
-7.73e+04	2.19e-01	-8.15e-01	3.08e+02	4.50e+02	1.10e+05

Table 3.44 Eigenvectors and eigenvalues for θ_1 and input $u_{1,1}$, combination system.

	Eigenvalue					
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
	-7.25e+05	6.61e+05	1.52e+05	9.38e+04	-3.28e+04	1.72e+03
	Eigenvector					
Axis	X_1	X_2	X_3	X_4	X_5	X_6
ε_t	-0.3786	0.8203	-0.0001	0.4257	-0.0003	-0.0000
R	0.8749	0.1659	-0.0000	0.4521	0.0022	-0.0000
L_m	-0.3020	-0.5469	-0.0001	0.7791	0.0019	-0.0000
k_c	0.0008	-0.0005	-0.8790	0.0010	-0.4764	0.0069
ε_c	0.0013	-0.0008	0.4767	0.0021	-0.8788	0.0037
α	-0.0102	-0.0227	0.0000	-0.0862	-0.0002	0.0000
S_2	0.0000	-0.0000	-0.0050	0.0000	-0.0135	-0.6487
S_1	0.0000	-0.0000	-0.0041	-0.0000	-0.0155	-0.5327
P_0	0.0000	-0.0000	-0.0021	-0.0000	-0.0192	0.5435
$B_t(0)$	0.0005	0.0002	-0.0000	0.0060	-0.0001	-0.0000
T_0	-0.0005	-0.0002	0.0000	-0.0061	0.0001	0.0000
	0.0000	-0.0000	-0.0000	0.0000	-0.0000	-0.0003
	Eigenvalue					
	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}
	-7.15e+02	-6.25e+02	4.54e+02	-2.93e+02	2.80e+02	8.76e-09
	Eigenvector					
Axis	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
ε_t	-0.0000	0.0058	0.0116	-0.0000	-0.0501	-0.0000
R	-0.0001	0.0065	0.0111	0.0000	-0.0504	-0.0000
L_m	-0.0001	0.0078	0.0100	0.0000	-0.0507	-0.0000
k_c	0.0173	0.0001	-0.0001	-0.0034	-0.0000	0.0000
ε_c	0.0213	0.0002	-0.0001	-0.0032	-0.0000	0.0000
α	-0.0001	0.0582	0.2687	-0.0000	-0.9573	-0.0000
S_2	-0.2213	-0.0002	0.0000	0.7280	0.0000	0.0003
S_1	-0.5508	-0.0006	0.0000	-0.6424	0.0000	0.0003
P_0	-0.8043	-0.0009	0.0001	0.2395	0.0000	-0.0000
$B_t(0)$	0.0011	-0.9944	-0.0684	0.0000	-0.0802	-0.0000
N(0)	-0.0002	0.0873	-0.9606	0.0000	-0.2637	-0.0000
T_0	-0.0002	-0.0000	0.0000	-0.0000	0.0000	-1.0000

Step 3. Generate Λ parameter ranking.

Based on Table 3.44, the Λ values and associated parameter ranking for this *a priori* setting are shown in Table 3.45.

Table 3.45 Λ_i values and their associated ranking based upon Hessian eigenstructure.

Ranking	Parameter	Submodel	Λ
1	ε_t	tree	8.57e+05
2	R	tree	7.86e+05
3	L_m	tree	6.54e+05
4	k_c	crop	1.50e+05
5	ε_c	crop	1.03e+05
6	α	tree	3.09e+04
7	S_2	crop	2.69e+03
8	S_1	crop	2.64e+03
9	P_0	crop	2.53e+03
10	$B_t(0)$	tree	1.77e+03
11	N(0)	tree	1.67e+03
12	T_0	crop	1.44e+00

This leads to the following $\chi_{1,k}$ matrix of candidate parameter vectors for estimation (Table 3.46).

Table 3.46 $\chi_{1,k}$ candidates for reduced-order parameter estimation for combination model.

Par.	Sub-model	$\chi_{1,1}$	$\chi_{1,2}$	$\chi_{1,3}$	$\chi_{1,4}$	$\chi_{1,5}$	$\chi_{1,6}$	$\chi_{1,7}$	$\chi_{1,8}$	$\chi_{1,9}$	$\chi_{1,10}$	$\chi_{1,11}$	$\chi_{1,12}$
ε_t	tree	X	X	X	X	X	X	X	X	X	X	X	X
R	tree		X	X	X	X	X	X	X	X	X	X	X
L_m	tree			X	X	X	X	X	X	X	X	X	X
k_c	crop				X	X	X	X	X	X	X	X	X
ε_c	crop					X	X	X	X	X	X	X	X
α	tree						X	X	X	X	X	X	X
S_2	crop							X	X	X	X	X	X
S_1	crop								X	X	X	X	X
P_0	crop									X	X	X	X
$B_t(0)$	tree										X	X	X
N(0)	tree											X	X
T_0	crop												X

Step 5 (prior to Step 4). Check goodness-of-fit of y_0 and $\hat{y}_j$.

Because one can be certain that optimizing only tree parameters will not suffice for meeting the fit criterion for the crop model, it is certain that the first three parameter sets ($\chi_{1,1}$, $\chi_{1,2}$, $\chi_{1,3}$) will be inadequate. Rather than test each $\chi_{1,k}$ and consistent identifiability, the optimization algorithm was run from θ_1 for the first seven $\chi_{1,k}$ and test the model fit for each. From there, a $\chi_{1,k}$ can be selected with good fit, followed by the consistent identifiability test, then validation of the estimated parameter set. Figure 3.31 shows the NRMSE for the first seven $\chi_{1,k}$. Both outputs meet the >99% criterion for $\chi_{1,5}$–$\chi_{1,7}$. Since the reduced-order set $\chi_{1,7}$ gives a bigger margin above 99%, we proceed with this optimized parameter set.

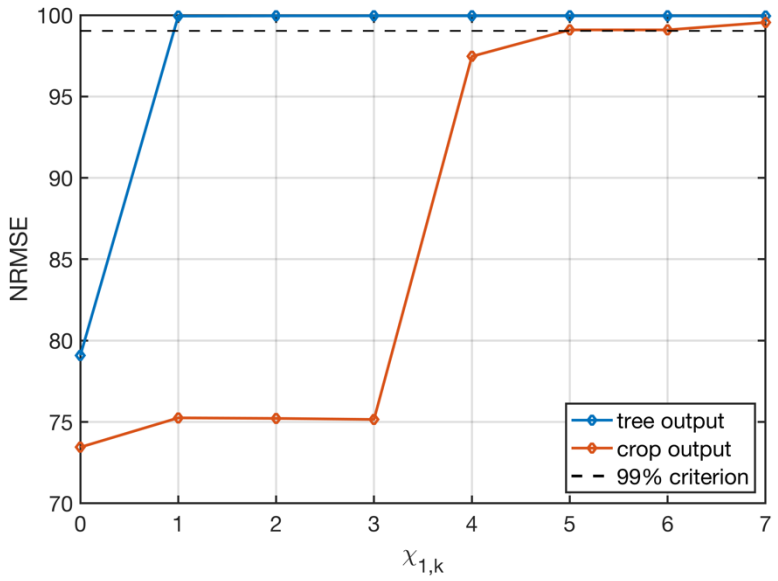

Figure 3.31 NRMSE values for unoptimized parameters (shown at 0) and the first seven $\chi_{1,k}$ optimized parameter sets. Both outputs meet the >99% criterion for $\chi_{1,5}$–$\chi_{1,7}$.

Step 4. Confirm consistent identifiability.

$\chi_{1,7}$ has seven estimated parameters, therefore, consistent identifiability must be tested from 2^7 or 128 *a priori* ±5% variations of the parameter estimated parameter values. All estimates converge to the same location $\hat{\hat{\theta}}_1$ within 4.64e–3%, as shown in Table 3.47. This confirms consistent identifiability for $\chi_{1,7}$.

Table 3.47 Step 4 for the first three candidate parameter sets.

	Optimized on $\chi_{1,7}$	*A priori* values	Actual
	$\hat{\hat{\theta}}_1$ (for all 128 a priori sets with ±5% values)	θ_1	θ_0
Max % param. diff.	4.64e-03%	–	–
ε_t	1.3860	1.3382	1.4086
α	9.0478e-05	9.5000e-05	0.0001
L_m	495.86	475.00	500
$B_t(0)$	95	95	100
N(0)	0.59137	0.59137	0.6225
R	189840	190000	200000
ε_c	0.30055	0.28500	0.3
T_0	0.475	0.475	0.5
P_0	0.76	0.76	0.8
S_1	1520	1520	1600
S_2	2060.7	1748	1840
k_c	0.41551	0.38	0.4
NRMSE tree	99.96%	79.09%	100%
NRMSE crop	99.51%	73.45%	100%

Step 6. Validation.

Validation with other four other inputs from L_1 is successful (Table 3.48).

Table 3.48 Fit measures with validation inputs from L_1 for $\hat{\hat{\theta}}_1$ optimized on $\chi_{1,7}$.

		Tree model output		Crop model output	
Input	**Use**	**NRMSE**	**Cost gm^2 tree^{-2}**	**NRMSE**	**Cost gm^2 m^{-2}**
$u_{1,1}$	Optimization	99.96%	5.3672e+04	99.51%	4.5973e-01
$u_{1,2}$	Validation	99.93%	9.2455e+04	99.50%	5.0933e-01
$u_{1,3}$	Validation	99.96%	3.4066e+04	99.41%	7.7433e-01
$u_{1,4}$	Validation	99.81%	5.1725e+05	99.09%	9.7463e-01
$u_{1,5}$	Validation	99.77%	6.6030e+05	99.06%	9.8339e-01

3.4.3.1 Simulations to show robustness

Once again, one may wonder how the $\hat{\hat{\theta}}_1$ parameterization optimized on $u_{1,1}$ performs for inputs from other locations. Table 3.49 shows that the procedure leads to parameters that are applicable to many different locations for the combination tree-crop growth model (in the water non-limiting case). In only one case (crop output for u_{41}), does the NRMSE value fall below 99%.

Table 3.49 Fit measures with validation inputs from nine locations other than the location used for optimization. In only one case does the fit measure drop below 99% (to 98.09%), indicating the parameterization $\hat{\hat{\theta}}_1$ optimized on $\chi_{1,7}$ is quite robust across locations.

		Tree model output		Crop model output	
Input	**Use**	**NRMSE**	**Cost gm^2 tree^{-2}**	**NRMSE**	**Cost gm^2 m^{-2}**
$u_{1,1}$	Optimization	99.95%	5.3672e+04	99.55%	4.5973e-01
$u_{2,1}$	Validation	99.73%	8.5818e+05	99.43%	5.3831e-01
$u_{3,1}$	Validation	99.32%	6.2902e+07	99.33%	3.7636e+00
$u_{4,1}$	Validation	99.68%	4.3511e+06	98.09%	1.0447e+01
$u_{5,1}$	Validation	99.34%	5.2171e+07	99.12%	5.2142e+00
$u_{6,1}$	Validation	99.84%	6.9812e+05	99.42%	1.0481e+00
$u_{7,1}$	Validation	99.95%	5.6831e+04	99.58%	3.4997e-01
$u_{8,1}$	Validation	99.76%	2.0029e+06	99.31%	1.5489e+00
$u_{9,1}$	Validation	99.55%	1.2413e+07	99.30%	2.4371e+00
$u_{10,1}$	Validation	99.60%	8.4069e+06	99.32%	1.8685e+00

3.4.3.2 Ranking by Λ for all 4096 θ_j for combination model

As a test of robustness of the Λ ranking, the Λ_i are calculated for all parameters and all 4096 θ_j. Table 3.50 shows the ranking of each parameter across all parameter sets. ε_t is the highest ranked parameter for all θ_j. This makes sense, as ε_t is most important for the tree model, and an accurate tree model is very important for a good fit of the crop model. Also of note, the top six parameters are the same for nearly all θ_j (with k_c and $B_t(0)$ switching places 5 times), suggesting some robustness of parameter choice for the highest ranking parameters. The rankings of the next five lower ranked parameters are mixed. The lowest ranked parameter is the same for all θ_j.

Figure 3.32 and Figure 3.33 show graphically the Λ_i values for the top eight ranked parameters (according to analysis of θ_1). These figures illustrate how Λ_i values cross over each other from θ_j to θ_j.

Table 3.50 Occurrence of parameter ranking for all 4096 *a priori* estimates θ_j.

		Ranking											
Par.	**Sub-model**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
ε_t	tree	4096	0	0	0	0	0	0	0	0	0	0	0
R	tree	0	3967	61	28	40	0	0	0	0	0	0	0
L_m	tree	0	0	3925	65	58	48	0	0	0	0	0	0
k_c	crop	0	113	42	2560	914	462	0	5	0	0	0	0
ε_c	crop	0	0	60	84	2375	1577	0	0	0	0	0	0
α	tree	0	16	8	1359	709	2004	0	0	0	0	0	0
S_2	crop	0	0	0	0	0	0	481	386	1253	1285	691	0
S_1	crop	0	0	0	0	0	0	892	534	1793	626	251	0
P_0	crop	0	0	0	0	0	0	110	528	718	822	1918	0
$B_t(0)$	tree	0	0	0	0	0	5	2283	476	162	868	302	0
N(0)	tree	0	0	0	0	0	0	330	2167	170	495	934	0
T_0	crop	0	0	0	0	0	0	0	0	0	0	0	4096

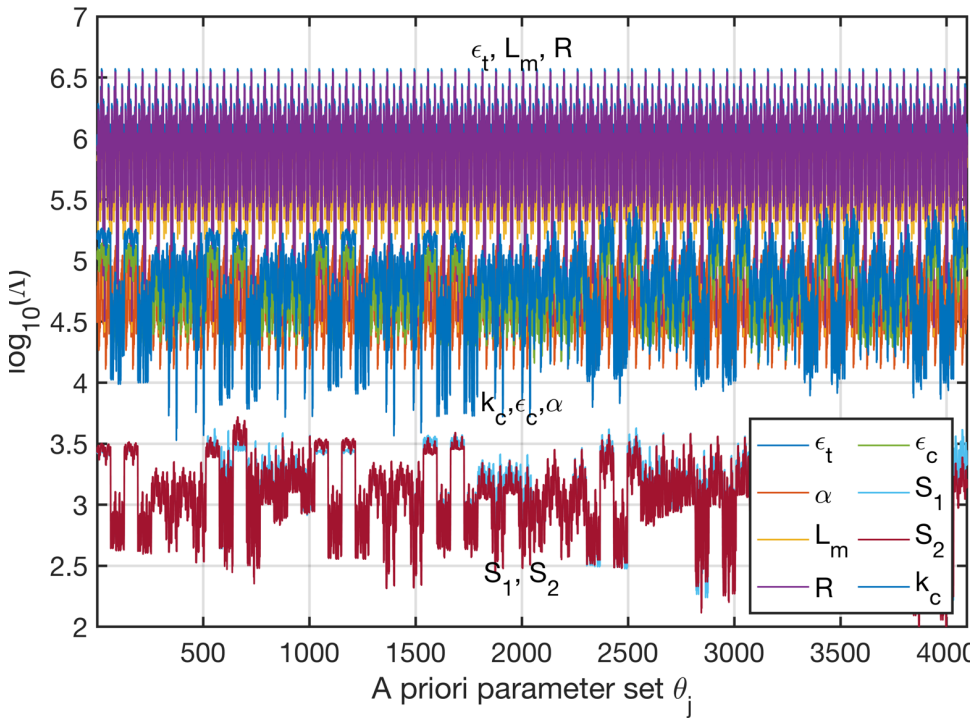

Figure 3.32 Λ_i for highest ranked eight parameters of the combination model across 4096 *a priori* estimates θ_j in $\log_{10}$ scale.

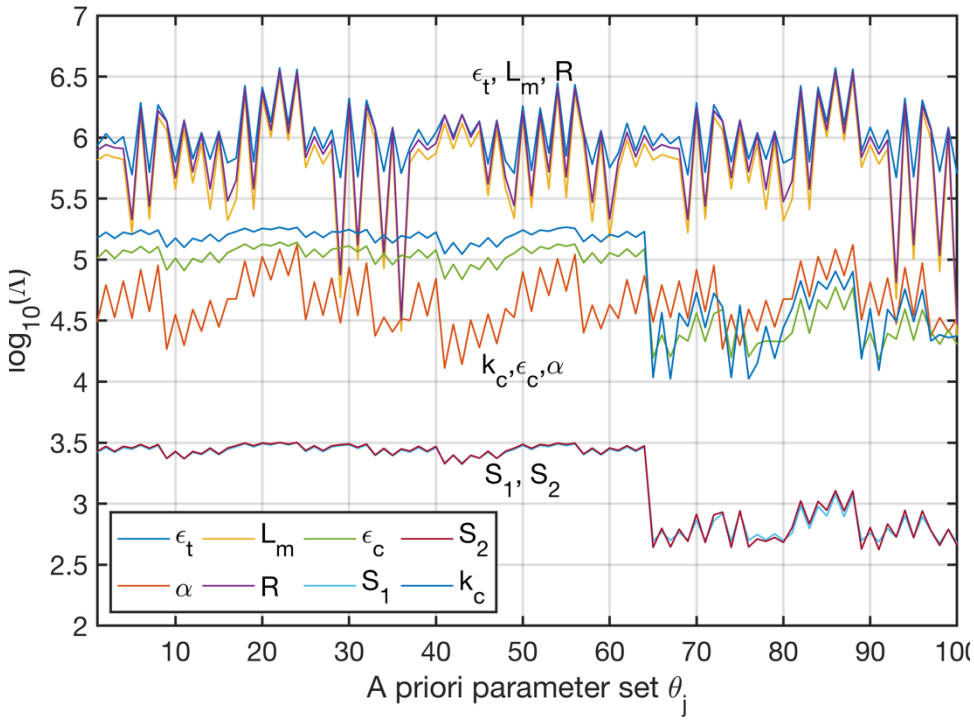

Figure 3.33 (Close up of Figure 3.32) Λ_i for highest ranked eight parameters of the combination model across 100 *a priori* estimates θ_j in $\log_{10}$ scale.

4 SIMULATION TOOLS AND DATA

This section covers the tools and data used in the model simulations. The reader may wish to skip over this chapter if the details of simulations are not of interest. Simulations were done in MATLAB 2017a and 2018a (MathWorks, https://www.mathworks.com/). Input data required to run the simulations was accessed through the CliPick portal (Palma 2017). These are described in detail below.

4.1 MATLAB implementation

4.1.1 Yield-SAFE model implementation and parameter estimation

Original sources for the Yield-SAFE model equations included van der Werf et al. (2007), Graves et al. (2010), and Keesman et al. (2011). The Microsoft Excel implementation by Burgess et al. (2014) generated the reference outputs used to check the MATLAB implementation. In other words, the Excel implementation was taken to be the actual system model. Matching the Excel outputs presented minor challenges, as the order of calculation in Excel is determined automatically by a rather opaque process. As Microsoft explains (n.d.), "Excel does not calculate cells in a fixed order, or by row or column. Instead, Excel dynamically determines the calculation sequence based on a list of all the formulas to calculate (the calculation chain) and the dependency information about each formula." Because of this, the time steps of model states $L_m(n)$ and $N(n)$ had to be adjusted in the MATLAB implementation to match the time indices executed in Excel.

MATLAB's Grey-Box Modeling tools that are part of the System Identification Toolbox™ were used for simulations and optimization (with exception of the

analytic Hessian as described in Section 4.1.3) (Ljung 2018, pp. 707–775). Primary MATLAB functions used in implementation of the HIROPE procedure in this dissertation are shown in Figure 4.1. For this purpose, the `idnlgrey` Nonlinear ODE (grey-box) object (https://www.mathworks.com/help/ident/ref/idnlgrey.html) was employed, which allows identification of selected model parameters and initial states. As described by Ljung (2018), "Identified Nonlinear Models represent nonlinear systems with coefficients that are identified using measured input/output data. You can specify initial values and constraints for the estimation of the coefficients." The `idnlgrey` model requires creating a function with a standardized format that contains the nonlinear model structure (equations), and a user-specified number of states, inputs, outputs, and parameters. A `idnlgrey` model is run with the `sim` function.

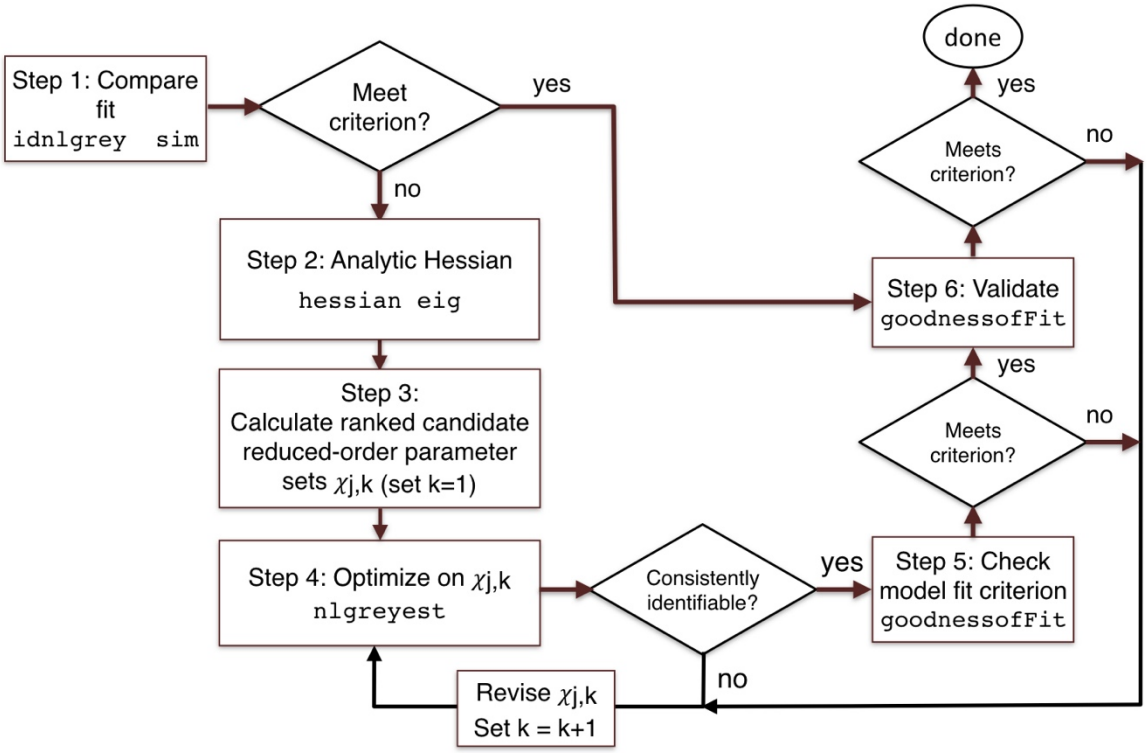

Figure 4.1 The HIROPE procedure with key MATLAB functions noted.

Once the `idnlgrey` model is set up, the `nlgreyest` (https://www.mathworks.com/help/ident/ref/nlgreyest.html) function is used to selectively estimate parameters and initial states of a nonlinear grey-box model defined by `idnlgrey`. There are a wide range of options for `idnlgrey` (https://www.mathworks.com/help/ident/ref/nlgreyestoptions.html), all of which were left at their default values with exception of the maximum number of iterations (`maxIterations`), which was increased from 20 to 100 for estimations of 3–6 parameters and 300 or more for estimation of more than 6 parameters. The value of `TolFun`, the termination criterion lower bound on the change in the value of the objective function during a step, was also kept at the default of 1e–5.

The tree model includes two parameters that are also initial states, $B_t(0)$ and $N(0)$. These were estimated by the `nlgreyest` algorithm by using `setinit` to "false" for the associated states. In order to have these two initial states also be parameters, two new initial states were added to the tree model structure, that acquire their estimated state values at each new optimization step and then are held constant during the simulation and fed into the growth equations in place of their respective parameters. This allowed for the simultaneous use of $B_t(0)$ and $N(0)$ as initial states and parameters in the optimization.

The default `lsqnonlin` option in `nlgreyest` was used for parameter estimation (except for the linear model examples, where the LMS algorithm was implemented in MATLAB code). The `lsqnonlin` function has as its default optimization algorithm the Trust-Region Reflective Newton algorithm. As stated in Ljung (2018), "The default trust-region-reflective algorithm is a subspace trust-

region method and is based on the interior-reflective Newton method described in [1] [Coleman and Li 1996] and [2] [Coleman and Li 1994]. Each iteration involves the approximate solution of a large linear system using the method of preconditioned conjugate gradients (PCG)." The trust region method has numerous implementations in many fields (Conn et al. 2000). The basic idea is to generate a model representative of the cost function within a region of the parameter space and then locate a point within this region that decreases the cost function. The size of this region is expanded or contracted depending on how well the model matches the cost function.

Regarding parameter scaling in the `idnlgrey` model, Ljung (2018, pp. 7-34) states, "When the model structure contains parameters with different orders of magnitude, try to scale the variables so that the parameters are all roughly the same magnitude." This confirms recommendations by others (Thacker 1989, Dennis and Schnabel 1996, Conn et al. 2000). However, running the simulations with and without the scaling sidecar variables s_{i} showed that optimization simulations took 2–3 times longer with scaling than without scaling (presumably due to the additional number of parameters), even though the same number of parameters were estimated in both. The optimized parameters converged to the same values with and without scaling in comparisons conducted with and without parameter scaling. As the simulation run times became prohibitively long, most optimizations were run without the scaling. No anomalies in outcomes were noticed.

4.1.2 OAT sensitivity analysis

One-at-time parameter sensitivity analysis was conducted simply by running the `idnlgrey` tree and crop models using the `sim` function. Each parameter of

interest was varied while holding the other parameters constant. The model fit metric was calculated using `goodnessofFit` with the `'NRMSE'` option (normalized root mean squared, as described in the text).

4.1.3 Analytic formulation of Hessian

The analytic Hessian is derived from the cost function expression. Manual differentiation of these expressions is possible, but they are complex and conversion to MATLAB equations would be intensely laborious and error-prone. Therefore, a twice continuously differentiable symbolic expression for the cost was first coded in MATLAB, using approximations for discontinuous functions in the growth equations where necessary. The Hessians of the symbolic cost function expressions (for a certain time step) were then derived using MATLAB's `hessian` function with respect to the parameters of interest. In order to calculate the average H value over a time period N, the symbolic Hessian is converted to a MATLAB function of the symbolic variables (parameters, other variables, and input, output, and state variables) via the `matlabFunction` function and the instantaneous Hessian terms are calculated and summed. Converting the Hessian to a MATLAB function has the advantage of speeding up the calculations by about two orders of magnitude as opposed to substituting values into the Hessian via the `subs` function at each step. The Hessian function is called in a `for` loop N times to calculate an average Hessian.

4.1.4 Eigenvalues and eigenvectors

The eigenstructure of the calculated Hessian `H` for the nonlinear models was determined using the `eig` function, with the option to return two matrices

$$[\mathrm{X}, D] = \mathrm{eig}(\mathrm{H}) \tag{3.30}$$

where D is the diagonal matrix of eigenvalues and X is a matrix whose columns are the corresponding right eigenvectors, such that $\mathrm{H}^* \mathrm{X} = \mathrm{X}^* D$. The eigenvectors of X (and corresponding columns of D) are not returned in any specific order. Since ordering by magnitude of the eigenvalues is of primary interest, the `sort` command was used to reorder D by magnitude followed by the same sort order of the corresponding columns of X.

4.1.5 Contours

Cost function contours across a 2-dimensional parameter space were generated first by calculating the cost over a 251 x 251 point grid (63,001 points). Cost function was calculated using the `sim` function for idnlgrey models or vector equations for the linear and logistic equations. The grid of cost values was converted to a contour plot using the `contour` function. A vector of contour levels is supplied as an argument to `contour`, then the function calculates the contour curves and plots them. As recommended by MATLAB (https://www.mathworks.com/help/matlab/data_analysis/convolution-filter-to-smooth-data.html), the cost data was smoothed using `conv2` with a 5 x 5 kernel. The smoothing was used especially to remove anomalies caused by the `contour` function particularly in flat regions near the curve minimum.

4.2 Generation of synthetic inputs

As real data of the duration and quality necessary for large scale tree growth simulations are lacking, climate model data were used to drive the Yield-SAFE simulations. There is increasing evidence that simulated data are a viable substitute for real data in forest growth modeling (Lamarque et al. 2011, Ailliot et al. 2015, Palma et al. 2018). Large-scale climate model data are available for

download from major national and international repositories (e.g., https://esgf-node.llnl.gov/projects/esgf-llnl/, http://www.cordex.org), however these data have many layers and are available in file formats that are not readily accessible for the purposes of forest modeling. Fortunately, the same project that currently develops the Yield-SAFE model also created a web portal named CliPick (Figure 4.2) to facilitate access to a number of climate datasets (Palma 2017), currently accessible at http://home.isa.utl.pt/~joaopalma/projects/agforward/clipick/. As noted by Palma (2017), programming was necessary in "Python, JavaScript and PHP, HTML, CSS and AJAX, in order to build a platform that could intuitively supply data in ASCII format."

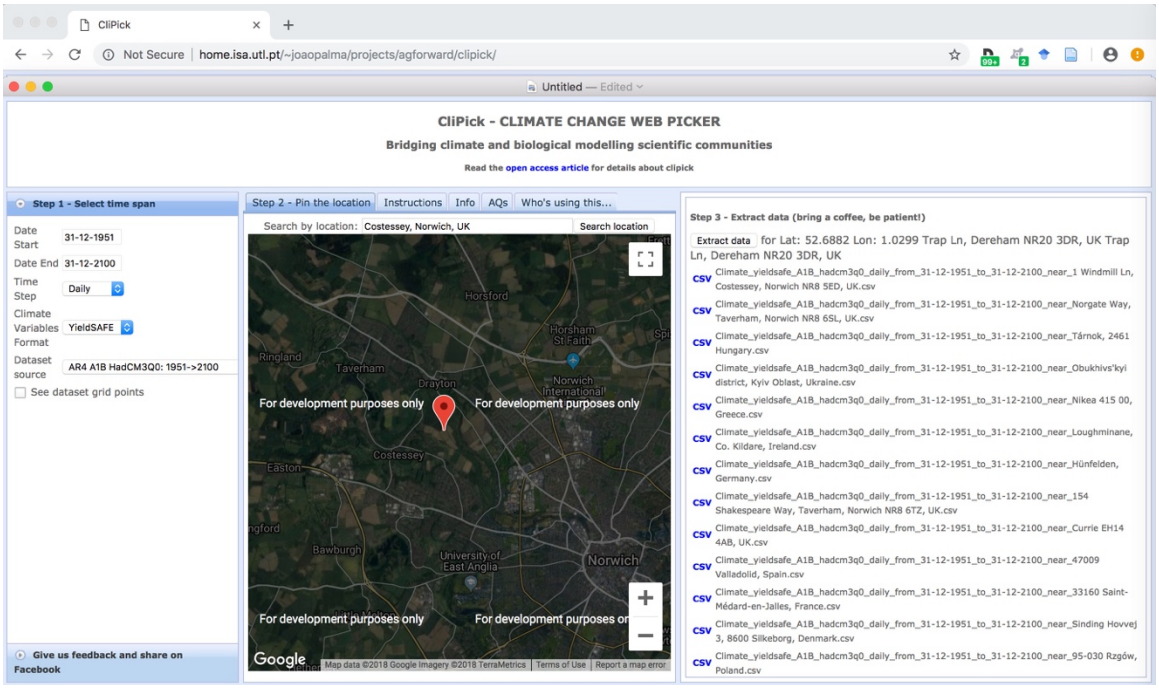

Figure 4.2 Screenshot of the CliPick interface. The left column allows user to select the time frame and desired dataset. The middle column selects the location either by text search or map interface. The right column contains the extracted dataset references for downloading.

Data were extracted via the CliPick tool for optimization and validation sequences. These include solar radiation and temperature, which were generated

with the climate model KNMI-RACMO22E (van Meijgaard et al. 2012). The full name of the dataset accessed through CliPick is "EUR-11_ICHEC-EC-EARTH_historical_r1i1p1_KNMI-RACMO22E_v1". For the simulations conducted here, temperature, solar radiation, and precipitation data from the period 1951–2005 were extracted. This length of data was sufficient to run the model for 20 years (i.e., 1951–70), leaving 35 years of other data for validation tests. The location for the initial analysis was "Lat: 52.6628 Lon: 1.2283 Costessey, Norwich, UK". For additional validation input sequences from a particular location, data from another IPCC climate change scenario at the same location, the "A1B - HadCM3Q0" dataset was used. Locations of data sources are given in Table 4.1.

Table 4.1. Input sequence location sources used for simulations. All data were imported through the CliPick portal.

Location symbol	Location description	Lat.	Long.	Data source	Start date	Length (# days)
L_1	Costessey, Norwich NR8 5ED, UK	52.6628	1.2283	a, b	31-12-51	7306
L_2	Currie EH14 4AB, UK	55.9058	-3.3275	a	31-12-51	7306
L_3	Nikea 415 00, Greece	39.4698	22.4530	a	31-12-51	7306
L_4	33160 Saint-Médard-en-Jalles, France	44.9058	-0.7758	a	31-12-51	7306
L_5	47009 Valladolid, Spain	41.7484	-4.8417	a	31-12-51	7306
L_6	Hünfelden, Germany	50.3127	8.1431	a	31-12-51	7306
L_7	8600 Silkeborg, Denmark	56.1941	9.5109	a	31-12-51	7306
L_8	95-030 Rzgów, Poland	51.6802	19.4996	a	31-12-51	7306
L_9	Tárnok, 2461 Hungary	47.3587	18.8407	a	31-12-51	7306
L_{10}	Obukhivskyi district, Kyivska oblast, Ukraine	50.1498	30.6447	a	31-12-51	7306

a: Hist KNMI-RACMO22E, b: A1B - HadCM3Q0

5 INVESTIGATION OF LOW-TECH DATA COLLECTION ISSUES EMPHASIZING ON-FARM PARTICIPATORY RESEARCH

5.1 Data is required for model validation and development

The high expense and long time frames required (measured in decades) for agroforestry field trials are primary motivations for developing reliable predictive models (Malézieux et al. 2009). Such models would give the ability to inexpensively test novel crop combinations and management regimes, instilling a measure of confidence before investing in new plantings. Additionally, dynamic models that include exogenous environmental growth drivers have the ability, at least theoretically, to predict outcomes within reasonable ranges of changing climate, which is important information for producers, policymakers, and insurers, among others.

In general, growth models consist of two components: system dynamics as described in mathematical equations and the equation parameters. System dynamics in agroforestry models such as WaNuCLAS, Hi-SAFE, and Yield-SAFE are constructed by carefully piecing together process-based dynamic blocks that have been well studied over many years or decades (Luedeling et al. 2016). In other words, the internal processes significant for growth are assumed to be well understood, and with this confidence the model structure is taken to be a white box model, a system model with well-known dynamics and parameters.

Yield-SAFE, arguably the only existing agroforestry model with viable predictive ability for production (tree and crop biomass), is parameterized based upon

monoculture models with a small portion of the parameters adjusted based on expert knowledge and the few available growth data (Graves et al. 2010). However, such an approach is wanting in two respects. First, it is limited to crops that have been modeled reliably in monocultures, which may not be the case for numerous tropical and temperate crops that are important for local economies, but are not considered major crops. Second, a small amount of data may be inadequate for reliable parameterization, especially when considering that plant growth dynamics change in multi-crop systems.

This second caution is pointed out in a recent review of tree-crop models (Luedeling et al. 2016) as motivation for developing new modeling methods, "Modeling approaches for capture of water, nutrients and light that produce reliable predictions in monocultures may not suffice in more complex situations, because they do not consider critical competitive and facilitative interactions between trees and crops." This recognition of complex interactive dynamics between crops indicates that long-term data sets are required to confirm (or put into question) underlying assumptions of process dynamics in any multi-crop grey box model built upon monoculture models and data.

5.2 Reduction of data requirements

Although as shown in previous chapters, an analysis of the Yield-SAFE model can be carried out in the absence of field data, data collection from field trials is essential for model validation and ultimately for reliable predictive ability. As Ljung (1998) states, "System identification begins and ends with real data. ... The result of the modeling process can be no better than what corresponds to the information contents in the data."

Fortunately, an I/O identification approach lends itself to reducing the data burden. The approach taken in previous chapters is based upon the premise that for I/O system identification all inputs are easily measurable and outputs are only quantities of interest in predicting and that are measurable. This means that rather than measuring outputs of the underlying biophysical processes such as number of branch nodes and leaf area in the case of Yield-SAFE, the previous analysis shows that only measures of tree and crop biomass need be collected in order to reliably parameterize the model (Table 5.1). Note once again that in this analysis framework, the parameters associated with the underlying process dynamics lose their physical interpretation.

Table 5.1. Data requirements for parameterization of all process sub-blocks in Yield-SAFE model compared with parameterization of I/O system identification approach to Yield-SAFE (based on van der Werf et al. 2007)

	Data required for parameterization of all individual Yield-SAFE sub-processes	**Data required for I/O system identification Yield-SAFE model**
Tree	number branch shoots*, leaf area index, radiation extinction coefficient*, water use*, relative effect of soil water potential*, biomass	biomass
Crop	leaf area index, water use*, relative effect of soil water potential*, solar radiation available, biomass	biomass
Soil	drainage of soil water below the potential rooting zone*, actual soil evaporation*, volumetric water content, radiation intercepted by trees, radiation interception by crop, soil water tension	soil water tension
Inputs	solar radiation, temperature, precipitation, irrigation	solar radiation, temperature, precipitation, irrigation

* not measurable or highly impractical to measure in field.

5.3 Justification for conducting agroforestry research on farms

Thirty years ago it was envisaged that process-based crop models would be developed for grower use. Whistler et al. (1986) state, "Some ... crop models are

being built with the ultimate goal of grower usage. This presents the exciting possibility of allowing precise, laboratory, growth chamber and/or field data to be used by the grower." The significant reduction in the number of measurements required for I/O modeling as compared with modeling of the underlying processes opens the door to moving research plots from research laboratories onto working farms.

In addition to opening the door to much larger data sets from realistic field conditions, participatory on-farm research has several compelling advantages for agroforestry. Because research must be carried out over long time periods (and therefore is costly to maintain), crop selection and configuration are tailored to the site and farmer preferences, and plots require frequent management attention. The prospect of carrying out research on farms greatly lowers the costs, while broadening the range of crops and environments covered. On farm studies also keeps researchers in touch with the needs of the end users, who are a heterogeneous group with interests that change over time (Landsberg 2003).

On-farm research in agroforestry systems such as alley cropping has been undertaken since the early years of agroforestry institutionalization in the mid-1970's (Atta-Krah and Francis 1987). However, recent technological advances further facilitate research activities outside of research institutions, including widespread use of smartphones (Dehnen-Schmutz et al. 2016), advances in small and inexpensive devices for field data collection and archiving (Aqeel-ur-Rehman et al. 2014), and widely available internet access for data upload/download. Finally, on-farm research fosters participatory synergies among producers, academic researchers, and even others such as policymakers, traditional leaders,

and entrepreneurs (Drinkwater et al. 2016). Together, these advances create a pathway for on-farm research for I/O modeling (Table 5.2).

Table 5.2. Pathway to on-farm research for I/O modeling.

Advantages over field station research	**Enabling technologies**	**Research framework**
Reduces research costs of field implementation Improves relevance to producers Enhances research design, implementation, and analysis	System identification model reduction Smartphone Automated environmental data collection Semi-automated data analysis software Cloud/Internet	Participatory, multi-disciplinary On-farm collaborative research networks Motivated by stakeholder needs Information flows both to/from producers

5.4 Scope of on-farm research in today's context

Thirty years ago, on-farm research was focused on application of agricultural technologies that had previously been developed and studied in research institutions. As stated by Atta-Krah and Francis (1987), "The purpose of on-farm trials is to obtain information about the performance of a technology under farm conditions." This could be said to be a top-down research model, where the foundational science takes place in controlled research stations and on-farm research was done to evaluate and demonstrate research recommendations.

Today, on-farm agricultural research garners wide acceptance for broad studies, including foundational research (Sooby 2001, Nielsen 2010, Schillinger 2011). Many researchers have shifted to a participatory model for on-farm research, where academic researchers and producers are equally important in project conceptualization, implementation, and analysis. As noted by Drinkwater et al. (2016), "Systems research teams can develop solutions that are applicable to real-

world situations, and this approach often provides a wider range of innovative solutions compared to a single-discipline approach."

A participatory, multi-disciplinary approach becomes particularly important as the subject of study becomes more complex (less reduced), where system behavior is increasingly affected by feedback within internal components and can take many years to reach a new steady state. This is the case with multistory agroforestry systems, which involve complex interactions between perennial crops, soil, and environment with the course of dynamic trajectories taking place over many years. In fact, many believe that multi-year studies are necessary even for annual crops, especially in transition from one management system to another (e.g., conventional to organic, till to no-till) (Drinkwater et al 2016).

5.5 Smartphone technologies

Technologies from the early space program such as photocells, anemometers, and data loggers with magnetic data storage helped enable a generation of models based on mechanistic process descriptions (Sinclair and Seligman 1996). The smartphone and inexpensive connected technologies for data acquisition promise a new period in farm-based research. Part of this shift from the laboratory to the farm involves the type of data required for model parameterization and validation. One goal of an input-output system identification approach is to require only data that is easy to measure using inexpensive and readily available equipment: a key strategic shift that could allow all needed data acquisition via smartphones.

There are currently over 2 billion smartphones in use and rates of use continue to rise rapidly worldwide (Poushter 2016). Because of their relatively low cost,

mobility, and computational abilities, smartphones are particularly applicable devices for use in on-farm research. Moreover, standard on-board physical sensors commonly found in smart phones such as camera, geo-referencing through GPS, microphone, and accelerometer give smartphones capabilities to collect many types of data without external measurement devices (Pongnumkul et al. 2015). Wireless connectivity via Wi-Fi, Bluetooth, NFC, and wireless cell phone networks, also give smartphones unprecedented capabilities to both capture data from external devices and send data to devices or the internet.

In the context of data collection in I/O system identification (Table 5.1), measurements of tree and crop biomass, soil water tension, and environmental inputs (solar radiation, temperature, and precipitation) are required. Biomass may be measured via tree size (e.g. height, DBH) which, depending on the application, may use a smartphone as an inclinometer to determine height. One could also envision a dedicated smartphone application that estimates biomass based on an image of a plant canopy taken from a reference distance (see e.g., Vastaranta 2015). Soil water tension and weather data would be collected by external devices, which transfer their data to the smartphone. All collected data is uploaded automatically from the smartphone to a central data collection web site.

In addition to I/O data collection, smartphones may also be used to collect other data relevant to a particular study. For example, the smartphone accelerometer may be used to track the activity of the person carrying it, allowing logging of labor inputs (Pongnumkul et al. 2015). Additional recommendations about plant health (using image analysis), irrigation (based on soil water tension), and

pruning (based on biomass estimates) could also conceivably be included as spinoffs of the smartphone-based research.

Because the research will take place over many years, it is crucial that data collection methods remain consistent. As recommended by Drinkwater et al. (2016), "For long-term studies, establish methods for archiving research protocols, samples and data. Develop a master document to archive field notes and descriptions of weather, field operations and sample collection, along with any observations of unique factors that may have influenced system performance. Finally, have a mechanism for maintaining continuity of management regimes, sample collection, analytical techniques and storage methods." Managing data collection through smartphone applications can help ensure that consistency is achieved.

Participatory on-farm research networks allow producers and researchers to collaborate in mutually beneficial ways. Murrell (2013) identified nine such networks in the United States. These networks provide data analysis and technical advice to members (and in some cases the general public). This is one promising model for providing feedback to producers based on data analysis from groups of participating growers, which can also be delivered via a smartphone.

5.6 Potential research questions to address through a participatory framework

In addition to collecting data for I/O system identification, larger research hypotheses may be simultaneously addressed when working with a multi-disciplinary team (Table 5.3). For example, by growing the same crops in multistory planting configurations in different locations, one can begin to address

the question of whether model parameters are indeed constant within a range of environmental conditions. Land Equivalent Ratio, a measure of total productivity of multi-crop systems as compared with monocultures of the same crops (Mead and Willey 1980), could be studied by including monoculture plots of the species used in the multistory configuration. As a complement to measuring productivity (in biomass), tracking of inputs such as labor, equipment hours, and materials would allow calculation of the cost of production and net present value estimates. Finally, a range of planting densities and management regimes could be integrated into experimental design for study of optimizing outcomes. All of these experimental treatments would be developed collaboratively by researchers and producers (Drinkwater et al. 2016).

Table 5.3. Potential research areas that could be explored in parallel to I/O modeling data collection

Multi-locational on-farm yield trials (Test if model parameters are consistent across environments)
Multistory plots of different densities
Multistory plots with different management regimes, such as pruning and irrigation
Cost of production, Net Present Value
Land Equivalent Ratio (multistory and monoculture plots)

6 CONCLUSIONS AND OPEN ISSUES

Reduced-order parameter estimation based on the Hessian of the cost function in this early investigation uses foundational systems theory, straightforward mathematics, and off-the-shelf simulation and optimization software (MATLAB), and is by no means a state-of-the-art implementation. Rather, the research presented here is meant to illustrate the value of *a priori* investigation of parameter identifiability through the conceptual framework of system identification. This small step forward may pique the interest of those challenged by parameter identification in plant growth models. The potential of a systematic approach to determining the importance of parameters in a complex model illuminates an area of model development that can be shrouded in guesswork and trial-and-error. In this sense, this project may catalyze new perspectives in *a priori* model analysis.

Initial tests of the HIROPE procedure on the Yield-SAFE model show promise in the 12-parameter, non-water limiting tree-crop combination model. Much more investigation is needed to test the procedure on more complex implementations of Yield-SAFE (with other connecting submodels included) and with other plant growth models before any definitive conclusions can be made about the applicability of HIROPE. Further refinement of the procedure will be needed. For example, as model complexity increases, it may be advantageous to refine the parameter ranking metric based on topological analysis of the cost function surface. Additionally, the definition of the ranking metric Λ used in this dissertation proved to be an effective measure for relative parameter influence on the cost function, however, this formulation deserves further study, especially as the dimensionality of the estimation problem increases.

The procedure proposed here is only a small part of model development (although it does involve a considerable amount of work). Real data are required to validate model predictions and give a model credibility. Even more important, one should know which data are needed for model parameterization before undertaking field study. Analysis such as that given by HIROPE not only suggests which parameters are most important for estimation, but in the process of validation one may generate indications of which data are needed for parameter identification. Therefore a procedure such as HIROPE conducted *a priori* can influence experimental design.

By shifting away from the need to measure internal model signals that are difficult to obtain, input-output identification relies only on easily measurable output data. This philosophy of model building opens doors to study of plant growth systems such as agroforestry on a broader scale in participatory, on-farm research. Engagement with traditional farmers with intensive experience with agroforestry and numerous crops may enhance experimental design and shed light on model parameterization for novel crops that have been little studied in a scientific context. Such partnerships could eventually lead to better understanding of the dynamics of complex multi-crop agriculture based on data from more environments and crop combinations. The ubiquity of smartphone technology and internet connectivity combined with inexpensive instrumentation for measurements and allows field experiments to generate more useful information at lower cost, putting this modelling strategy within reach for the first time in history.

6.1 Open issues

6.1.1 Determination of initial parameter estimates

For a real system, the parameters are unknown. Because of model nonlinearities and noisy measurements, initial estimates too far away from the global minimum may lead to stalling or attraction to local minima. Therefore, one would like some assurance that the initial parameter estimates are within a region of attraction to the global minimum. There are three primary methods used for parameter initialization (Parillo and Ljung 2003)

1. Use the physical interpretation of the parameters to estimate parameters based upon experience and expert knowledge. This method combined with setting constraints on parameter values was used in Yield-SAFE parameter optimization (Keesman et al. 2009). Although this is an attractive path forward, the I/O approach allows parameters to stray from their physical values in order to fit the data, therefore certain model parameters may lose their physical interpretation. Additionally, as model complexity increases nonlinearities undermine confidence that physically interpretable parameters lie within the global minimum's region of attraction.
2. A brute force multiple-run approach of initializing parameters on a grid or randomly throughout the reasonable parameter space. Such methods have been used with process-based tree growth models (e.g., Mäkelä et al. 2007). Parillo and Ljung (2003) show examples of identifying parameters of stable low-order models in the noise-free case with a 0% success rate over 100 random initializations, suggesting that even brute force methods might not be successful.

3. Use algebraic tools to rearrange the model. Parillo and Ljung (2003) use the example of Ljung and Glad's (1994) method of rearranging any globally identifiable structure to a linear regression, which could be used as an initial parameter estimate. The computational costs are, however, for high dimensional systems are considered possibly prohibitive (Ljung 2010). To address this limitation, Mercère et al. (2014) use a linear time-invariant black box model and non-iterative identification algorithm to arrive an accurately parameterized black-box model, then transforming the black box model into a state-space representation with parameters corresponding to the original model that can be used as initial parameter estimates.

6.1.2 Are model parameters time-varying?

It is likely that growth dynamics vary with time (e.g., with maturity of the trees and other crops) across different planting configurations, especially with multiple species and management routines. Young (1983, p. 70) comments on the ambiguities inherent in models of natural systems, "The size and complexity of many natural systems, such as those encountered in environmental and economic research, are such that the mechanisms governing the change in the observed system variables and their interrelationships are rarely fully understood a priori." For example, when two different crop species are placed in close proximity, interactions between them might alter their growth rates in fundamental ways though competition and facilitation dynamics (Pugnaire et al. 1996, Callaway and Walker 1997, Cavard et al. 2011, Fichtner et al. 2017). Regarding balance between model utility and capturing complex biological dynamics, Monteith (1996) states, "... simulation models [are] composed of dozens, even hundreds of

algorithms, each containing a set of empirically determined constants. A major but usually unstated assumption is that such constants can be transferred from site to site, from season to season, and sometimes even from species to species." In the model, then, it is unknown which submodel parameterizations stay constant and which vary, and further, how varying parameters may influence the complex nonlinear model dynamics. Models at a higher level on the hierarchy of systems detail, such as sigmoid models, may be less prone to this problem (Zeide 1996, Young 2012).

6.1.3 Field data collection strategies and experimental design

As this model analysis approach is based upon simulated data and is meant as a precursor to field trials and data collection, it is natural to extend the procedure to experimental design. This could include optimizing field experiments to maximize the amount of useful information gained, while minimizing associated cost (Hengl et al. 2007). This could be done by investigation of the sampling interval required at which data are needed (Walter 2012). Integration of available remote sensing data (collected remotely by satellites, aircraft, or land towers) should also be considered (Zarco-Tejada et al. 2016). Another design consideration is the location of studies—regions with richer variation in weather at an appropriate time scale might yield better data for system identification, and therefore should be prioritized.

7 REFERENCES

Abusam, A., K. J. Keesman, G. Van Straten, H. Spanjers, and K. Meinema. 2001. Sensitivity Analysis in Oxidation Ditch Modelling: The Effect of Variations in Stoichiometric, Kinetic and Operating Parameters on the Performance Indices. *Journal of Chemical Technology & Biotechnology: International Research in Process, Environmental & Clean Technology* 76 (4): 430–438.

Affholder, F., P. Tittonell, M. Corbeels, S. Roux, N. Motisi, P. Tixier, and J. Wery. 2012. Ad Hoc Modeling in Agronomy: What Have We Learned in the Last 15 Years? *Agronomy Journal* 104 (3): 735. doi:10.2134/agronj2011.0376.

AFRI (Agriculture and Food Research Initiative). 2016. Competitive Grants Program, Foundational Program FY 2016 Request for Applications (RFA). https://nifa.usda.gov/sites/default/files/rfa/FY%202016%20AFRI%20Foundational%20Program%20RFA_revised%20May%2027%202016.pdf

Ailliot, P., D. Allard, V. Monbet, and P. Naveau. 2015. Stochastic Weather Generators: An Overview of Weather Type Models. http://ciam.inra.fr/biosp/sites/ciam.inra.fr.biosp/files/WeatherType_SWG_Review_AAMN.pdf.

Aqeel-ur-Rehman, A. Z. A., Noman I., and Z. A. Shaikh. 2014. A Review of Wireless Sensors and Networks' Applications in Agriculture. *Computer Standards & Interfaces* 36 (2): 263–70. doi:10.1016/j.csi.2011.03.004.

Aster, R. C., Brian B., and C. H. Thurber. 2011. *Parameter Estimation and Inverse Problems*. Vol. 90. Academic Press.

Åström, K. J. 1967. Computer Control of a Paper Machine: An Application of Linear Stochastic Control Theory. *IBM Journal of Research and Development,* 11(4), 389-405. DOI: 10.1147/rd.114.0389

Åström, K. J., and P. Eykhoff. 1971. System Identification—a Survey. *Automatica* 7 (2): 123–62.

Atta-Krah, A. N., and P. A. Francis. 1987. The Role of on-Farm Trails in the Evaluation of Composite Technologies: The Case of Alley Farming in Southern Nigeria. *Agricultural Systems* 23 (2): 133–52. doi:10.1016/0308-521X(87)90091-6.

Audoly, S., Bellu, G., L. D'Angio, M. P. Saccomani, and C. Cobelli. 2001. Global identifiability of nonlinear models of biological systems. *IEEE Transactions on Biomedical Engineering*, 48(1), 55–65. https://doi.org/10.1109/10.900248

Balakrishnan, A.V., and V. Peterka. 1969. Identification in Automatic Control Systems. *Automatica* 5 (6): 817–29. doi:10.1016/0005-1098(69)90095-8.

Bashir, Omar, K. Willcox, O. Ghattas, B. van Bloemen Waanders, and J. Hill. 2008. Hessian-Based Model Reduction for Large-Scale Systems with Initial-Condition Inputs. *International Journal for Numerical Methods in Engineering* 73 (6): 844–868.

Batho, A., and O. Garcia. 2006. De Perthuis and the origins of site index: a historical note. *FBMIS*, 1, 1–10.

Beck, M. B. 1983. Uncertainty, system identification, and the prediction of water quality. *Uncertainty and Forecasting of Water Quality*, 3–68.

Bellman, R., and K. J. Åström. 1970. On structural identifiability. *Mathematical Biosciences*, 7(3), 329–339. https://doi.org/10.1016/0025-5564(70)90132-X

Berkes, F., J. Colding, and C. Folke. 2000. Rediscovery of traditional ecological knowledge as adaptive management. *Ecological Applications*, 10(5), 1251–1262.

Beven, K. J. 2002. Uncertainty and the detection of structural change in models of environmental systems. In: Beck MB (ed) Environmental foresight and models: a manifesto. Elsevier, The Netherlands, pp. 227–250.

Black, C., and C. Ong. 2000. Utilisation of light and water in tropical agriculture. *Agricultural and Forest Meteorology*, 104(1), 25–47. https://doi.org/10.1016/S0168-1923(00)00145-3

Blumer, A., A. Ehrenfeucht, D. Haussler, and M. K. Warmuth. 1987. Occam's Razor. *Information Processing Letters* 24 (6): 377–80. https://doi.org/10.1016/0020-0190(87)90114-1.

Boote, K. J., J. W. Jones, and N. B. Pickering. 1996. Potential Uses and Limitations of Crop Models. *Agronomy Journal* 88 (5): 704. doi:10.2134/agronj1996.00021962008800050005x.

Brisson, N., C. Gary, E. Justes, R. Roche, B. Mary, D. Ripoche, D. Zimmer, et al. 2003. An Overview of the Crop Model STICS. *European Journal of Agronomy* 18 (3): 309–332.

Buck-Sorlin, G. 2013. Process-Based Model. In *Encyclopedia of Systems Biology*, edited by W. Dubitzky, O. Wolkenhauer, K.-H. Cho, and H. Yokota, 1755–1755. Springer, New York. doi:10.1007/978-1-4419-9863-7_1545.

Burgess, P., A. Graves, and M. Upson. 2014. Yield-SAFE in Excel. Cranfield University, Cranfield, Bedfordshire, UK.

Byrd, R. H., P. Lu, J. Nocedal, and C. Zhu. 1995. A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM Journal on Scientific Computing* 16 (5): 1190–1208. doi: 10.1137/0916069

Callaway, R. M., and L. R. Walker. 1997. Competition and facilitation: a synthetic approach to interactions in plant communities. *Ecology*, *78*(7), 1958–1965.

Campbell, G. S., and J. M. Norman. 2012. *An introduction to environmental biophysics*. Springer Science & Business Media.

Cavard, X., Y. Bergeron, H. Y. H. Chen, D. Paré, J. Laganière, and B. Brassard. 2011. Competition and facilitation between tree species change with stand development. *Oikos*, *120*(11), 1683–1695. https://doi.org/10.1111/j.1600-0706.2011.19294.x

Coleman, T. F., and Y. Li. 1994. On the Convergence of Interior-Reflective Newton Methods for Nonlinear Minimization Subject to Bounds. *Mathematical Programming* 67 (1–3): 189–224.

Coleman, T. F., and Y. Li. 1996. An Interior Trust Region Approach for Nonlinear Minimization Subject to Bounds. *SIAM Journal on Optimization; Philadelphia* 6 (2): 28. http://dx.doi.org.proxy.library.cornell.edu/10.1137/0806023.

Conn, A., N. Gould, and P. Toint. 2000. *Trust Region Methods*. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9780898719857.

Cournède, P.-H., V. Letort, A. Mathieu, M. Z. Kang, S. Lemaire, S. Trevezas, F. Houllier, and P. De Reffye. 2011. Some Parameter Estimation Issues in Functional-Structural Plant Modelling. *Mathematical Modelling of Natural Phenomena* 6 (2): 133–59. doi:10.1051/mmnp/20116205.

Cox, G. M., J. M. Gibbons, A. T. A. Wood, J. Craigon, S. J. Ramsden, and N. M. J. Crout. 2006. Towards the Systematic Simplification of Mechanistic Models. *Ecological Modelling* 198 (1–2): 240–46. doi:10.1016/j.ecolmodel.2006.04.016.

Crout, N. M. J., D. Tarsitano, and A. T. Wood. 2009. Is My Model Too Complex? Evaluating Model Formulation Using Model Reduction. *Environmental Modelling & Software* 24 (1): 1–7. doi:10.1016/j.envsoft.2008.06.004.

de Reffye, P., E. Heuvelink, D. Barthélémy, and P. H. Cournède. 2008. Plant Growth Models A2 - Jørgensen, Sven Erik. In *Encyclopedia of Ecology*, edited by B. D. Fath, 2824–37. Oxford: Academic Press. http://www.sciencedirect.com/science/article/pii/B9780080454054002172.

Dehnen-Schmutz, K., G. L. Foster, L. Owen, and S. Persello. 2016. Exploring the Role of Smartphone Technology for Citizen Science in Agriculture. *Agronomy for Sustainable Development* 36 (2): 1–8.

Dennis Jr, J. E., and R. B. Schnabel. 1996. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Vol. 16. Siam.

Domingos, P. 1999. The Role of Occam's Razor in Knowledge Discovery. *Data Mining and Knowledge Discovery* 3 (4): 409–425.

Drinkwater, Laurie, Diana Friedman, and Louise E. Buck. 2016. *Systems Research for Agriculture: Innovative Solutions to Complex Challenges*. Brentwood, MD: Sustainable Agriculture Research and Education (SARE).

Dupraz, C., P. J. Burgess, A. Gavaland, A. R. Graves, F. Herzog, L. D. Incoll, N. Jackson, K. Keesman, G. Lawson, and I. Lecomte. 2005. *SAFE (Silvoarable Agroforestry for Europe) Final Report. SAFE Project (August 2001-January 2005).* Montpellier, France: INRA.

Elevitch, C.R., D.N. Mazaroli, and D. Ragone. 2018. Agroforestry Standards for Regenerative Agriculture. *Sustainability* 10 (9): 3337. https://doi.org/10.3390/su10093337.

Eykhoff, P. 1981. *Trends and Progress in System Identification: IFAC Series for Graduates, Research Workers & Practising Engineers*. Elsevier.

Fan, X.-R., M.-Z. Kang, E. Heuvelink, P. de Reffye, and B.-G. Hu. 2015. A knowledge-and-data-driven modeling approach for simulating plant growth: A case study on tomato growth. *Ecological Modelling*, 312, 363–373.

Fichtner, A., W. Härdtle, Y. Li, H. Bruelheide, M. Kunz, and G. von Oheimb. 2017. From competition to facilitation: how tree species respond to neighbourhood diversity. *Ecology Letters*, *20*(7), 892–900. https://doi.org/10.1111/ele.12786

Fontes, L., J.-D. Bontemps, H. Bugmann, M. Van Oijen, C. Gracia, K. Kramer, ... J. P. Skovsgaard. 2010. Models for supporting forest management in a changing environment. *Forest Systems*, *3*(4), 8–29.

Gadgil, M., F. Berkes, and C. Folke. 1993. Indigenous Knowledge for Biodiversity Conservation. *Ambio*, 151–156.

García, O. 2011. Dynamical implications of the variability representation in site-index modelling. *European Journal of Forest Research*, 130(4), 671–675.

Gerdin, M., T. Glad, and L. Ljung. 2007. Global Identifiability of Complex Models, Constructed from Simple Submodels. *Modeling, Estimation and Control*, 123–133.

Godfrey, K. R., and R. F. Brown. 1979. Practical aspects of the identification of process dynamics. *Transactions of the Institute of Measurement and Control*, *1*(2), 85–95. https://doi.org/10.1177/014233127900100204

Goudriaan, J., and H. H. Van Laar. 1994. *Modelling potential crop growth processes: textbook with exercises* (Vol. 2). Springer.

Graves, A. R., P. J. Burgess, J. Palma, K. J. Keesman, W. van der Werf, C. Dupraz, H. van Keulen, F. Herzog, and M. Mayus. 2010. Implementation and Calibration of the Parameter-Sparse Yield-SAFE Model to Predict Production and Land Equivalent Ratio in Mixed Tree and Crop Systems under Two Contrasting Production Situations in Europe. *Ecological Modelling* 221 (13–14): 1744–56. doi:10.1016/j.ecolmodel.2010.03.008.

Gutenkunst, R. N., J. J. Waterfall, F. P. Casey, K. S. Brown, C. R. Myers, and J. P. Sethna. 2007. Universally Sloppy Parameter Sensitivities in Systems Biology Models. *PLOS Computational Biology* 3 (10): e189. https://doi.org/10.1371/journal.pcbi.0030189.

Hengl, S., C. Kreutz, J. Timmer, and T. Maiwald. 2007. Data-based identifiability analysis of non-linear dynamical models. *Bioinformatics*, 23(19), 2612–2618. https://doi.org/10.1093/bioinformatics/btm382

Heuvelink, E., P. de Reffye. 2015. UVED Resource Plant Growth Architecture and Production Dynamics Preliminary Course: Plant & Crop models. CIRAD. http://greenlab.cirad.fr/GLUVED/html/P1_Prelim/Model/Model_PrelimCourse.pdf

Hillbrand, A., S. Borelli, M. Conigliaro, and A. Olivier, 2017. *Agroforestry for Landscape Restoration: Exploring the Potential of Agroforestry to Enhance the Sustainability and Resilience of Degraded Landscapes*; Food and Agriculture Organization of the United Nations: Rome, Italy.

Hinze, M., and S. Volkwein. 2005. Proper orthogonal decomposition surrogate models for nonlinear dynamical systems: Error estimates and suboptimal control. P. Benner, V. Mehrmann, and D. C. Sorensen, *Dimension reduction of large-scale systems*. Vol. 35. pp. 261–306. Springer-Verlag: Berlin.

Hornberger, G. M., and R. C. Spear. 1981. Approach to the Preliminary Analysis of Environmental Systems. *J. Environ. Mgmt.* 12 (1): 7–18.

IPCC. In press. Global warming of 1.5°C. Intergovernmental Panel on Climate Change. http://www.ipcc.ch/report/sr15/

Isaac, M. E., E. Dawoe, and K. Sieciechowicz. 2008. Assessing Local Knowledge Use in Agroforestry Management with Cognitive Maps. *Environmental Management*, 43(6), 1321–1329. https://doi.org/10.1007/s00267-008-9201-8

Jakeman, A. J., R. A. Letcher, and J. P. Norton. 2006. Ten Iterative Steps in Development and Evaluation of Environmental Models. *Environmental Modelling & Software* 21 (5): 602–14. doi:10.1016/j.envsoft.2006.01.004.

Jarvis, P. G. 1981. Plant water relations in models of tree growth. Studia Forestalia Suecia (Sweden). Retrieved from http://agris.fao.org/agris-search/search.do?recordID=SE8200906

Jat, M. L., J. C. Dagar, T. B. Sapkota, B. Govaerts, S. L. Ridaura, Y. S. Saharawat, R. K. Sharma, J. P. Tetarwal, R. K. Jat, H. Hobbs, et al. 2016. Chapter Three—Climate change and agriculture: Adaptation strategies and mitigation opportunities for food security in South Asia and Latin America. In *Advances in Agron*; Sparks, D. L., Ed.; Academic Press: Cambridge, Massachusetts, USA. Volume 137, pp. 127–235, doi:10.1016/bs.agron.2015.12.005.

Johnsen, K., L. Samuelson, R. Teskey, S. McNulty, and T. Fox. 2001. Process Models as Tools in Forestry Research and Management. *Forest Science* 47 (1): 2–8.

Johnson, Jr., C. R. 1988. *Lectures & Adaptive Parameter Estimation*. Prentice-Hall, Inc.

Jose, S. 2009. Agroforestry for ecosystem services and environmental benefits: An overview. *Agroforest. Syst. 76*, 1–10, doi:10.1007/s10457-009-9229-7.

Keesman, K. J. 1989. On the Dominance of Parameters in Structural Models of Ill-Defined Systems. *Applied Mathematics and Computation* 30 (2): 133–47. https://doi.org/10.1016/0096-3003(89)90147-1.

Keesman, K. J. 2011. *System identification: an introduction*. Springer.

Keesman, K. J., and R. Stappers. 2004. Nonlinear Set-Membership Estimation: A Support Vector Machine Approach. *Journal of Inverse and Ill-Posed Problems Jiip* 12 (1): 27–41.

Keesman, K. J., A. Graves, W. van der Werf, P. J. Burgess, J. Palma, C. Dupraz, and H. van Keulen. 2011. A System Identification Approach for Developing and Parameterising an Agroforestry System Model under Constrained Availability of Data. *Environmental Modelling & Software* 26 (12): 1540–53. doi:10.1016/j.envsoft.2011.07.020.

Kerschen, G., J. Golinval, A. F. Vakakis, and L. A. Bergman. 2005. The method of proper orthogonal decomposition for dynamical characterization and order reduction of mechanical systems: an overview. *Nonlinear Dynamics*, *41*(1), 147–169.

Korzukhin, M. D., M. T. Ter-Mikaelian, and R. G. Wagner. 1996. Process versus empirical models: which approach for forest ecosystem management? *Canadian Journal of Forest Research*, 26(5), 879–887. https://doi.org/10.1139/x26-096

Lamarque, J.-F., G. P. Kyle, M. Meinshausen, K. Riahi, S. J. Smith, D. P. van Vuuren, A. J. Conley, and F. Vitt. 2011. Global and Regional Evolution of Short-

Lived Radiatively-Active Gases and Aerosols in the Representative Concentration Pathways. *Climatic Change* 109 (1–2): 191.

Landsberg, J. 1981. The number and quality of the driving variables needed to model tree growth [environmental conditions]. *Studia Forestalia Suecia* (Sweden). No. 160. Retrieved from http://agris.fao.org/agris-search/search.do?recordID=SE19820831977

Landsberg, J. 2003a. Modelling forest ecosystems: state of the art, challenges, and future directions. *Canadian Journal of Forest Research*, 33(3), 385–397. https://doi.org/10.1139/x02-129

Landsberg, J. 2003b. Physiology in forest models: history and the future. *FBMIS*, 1, 49–63.

Landsberg, J., and P. Sands. 2011a. Modelling Tree Growth: Concepts and Review. Chapter 8. In *Terrestrial Ecology* (Volume 4, pp. 221–240). Elsevier. Retrieved from http://www.sciencedirect.com/science/article/pii/B9780123744609000081

Landsberg, J., and P. Sands. 2011b. Physiological ecology of forest production: principles, processes and models (Vol. 4). Academic Press.

Li, K.-C. 1992. On Principal Hessian Directions for Data Visualization and Dimension Reduction: Another Application of Stein's Lemma. *Journal of the American Statistical Association* 87 (420): 1025–1039.

Li, P., Q. D. Vu. 2013. Identification of parameter correlations for parameter estimation in dynamic biological models. *BMC Systems Biology*, *7*, 91. https://doi.org/10.1186/1752-0509-7-91

Linder, S. 1982. *Understanding and Predicting Tree Growth*. 160. http://pub.epsilon.slu.se/id/eprint/5287.

Ljung, L. 1996. Development of System Identification. Department of Electrical Engineering, Linköping University, Sweden. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.56.2331.

Ljung, L. 1999. *System Identification: Theory for the User*, Second Edition. Prentice Hall.

Ljung, L. 2010. Perspectives on System Identification. Annual Reviews in Control 34 (1): 1–12. doi:10.1016/j.arcontrol.2009.12.001.

Ljung, L. 2018. *System Identification Toolbox^{TM} User's Guide*. Natick, MA, USA: The MathWorks, Inc. https://www.mathworks.com/help/releases/R2018a/pdf_doc/ident/ident.pdf.

Ljung, L., and T. Glad. 1994a. *Modeling of Dynamic Systems*. PTR Prentice Hall Englewood Cliffs.

Ljung, L., and T. Glad. 1994b. On Global Identifiability for Arbitrary Model Parametrizations. *Automatica* 30 (2): 265–76. doi:10.1016/0005-1098(94)90029-9.

Luedeling, E., P. J. Smethurst, F. Baudron, J. Bayala, N. I. Huth, M. van Noordwijk, C. K. Ong, et al. 2016. Field-Scale Modeling of Tree–crop Interactions: Challenges and Development Needs. *Agricultural Systems* 142 (February): 51–69. doi:10.1016/j.agsy.2015.11.005.

Luo, Y., E. Weng, X. Wu, C. Gao, X. Zhou, and L. Zhang. 2009. Parameter Identifiability, Constraint, and Equifinality in Data Assimilation with Ecosystem Models. *Ecological Applications* 19 (3): 571–74.

Maclean, H., D. Dochain, G. Waters, M. Stasiak, M. Dixon, and D. Van Der Straeten. 2012. A Model Development Approach to Ensure Identifiability of a Simple Mass Balance Model for Photosynthesis and Respiration in a Plant Growth Chamber. *Ecological Modelling* 246 (November): 105–18. doi:10.1016/j.ecolmodel.2012.07.021.

Mäkelä, A., J. Landsberg, A. R. Ek, T. E. Burk, Mi. Ter-Mikaelian, G. I. Ågren, C. D. Oliver, and P. Puttonen. 2000. Process-Based Models for Forest Ecosystem Management: Current State of the Art and Challenges for Practical Implementation. *Tree Physiology* 20 (5–6): 289–298.

Mäkelä, A., M. Pulkkinen, P. Kolari, F. Lagergren, P. Berbigier, A., Lindroth, ... and P. Hari. 2008. Developing an empirical model of stand GPP with the LUE approach: analysis of eddy covariance data at five contrasting conifer sites in Europe. *Global Change Biology*, 14(1), 92–108. https://doi.org/10.1111/j.1365-2486.2007.01463.x

Malézieux, E., Y. Crozat, C. Dupraz, M. Laurans, D. Makowski, H. Ozier-Lafontaine, B. Rapidel, S. De Tourdonnet, and M. Valantin-Morison. 2009. Mixing Plant Species in Cropping Systems: Concepts, Tools and Models: A Review. In *Sustainable Agriculture*, 329–353. Springer.

Mason, E. G., R. W. Rose, L. S. Rosner. 2007. Time vs. light: a potentially useable light sum hybrid model to represent the juvenile growth of Douglas-fir subject to varying levels of competition. *Canadian Journal of Forest Research*, 37(4), 795–805.

MathWorks. 2018. R2018a Documentation. Goodness of fit between test and reference data - MATLAB goodnessOfFit. https://www.mathworks.com/help/ident/ref/goodnessoffit.html?searchHighlight=Normalized%20root%20mean%20square%20error&s_tid=doc_srchtitle [Accessed January 6, 2018].

Matthews, R. W., T. A. R. Jenkins, E. D. Mackie, E. D., and E. C. Dick. 2016. *Forest Yield: A handbook on forest growth and yield tables.* British forestry Forestry Commission, Edinburgh.

Mead, R., and R. W. Willey. 1980. The Concept of a 'land Equivalent Ratio' and Advantages in Yields from Intercropping. *Experimental Agriculture* 16 (3): 217–228. doi:http://dx.doi.org/10.1017/S0014479700010978.

Mendel, J. M. 1973. *Discrete Techniques of Parameter Estimation.* New York: Marcel Dekker.

Mercère, G., O. Prot, J. A. Ramos. 2014. Identification of parameterized gray-box state-space systems: From a black-box linear time-invariant representation to a structured one. *IEEE Transactions on Automatic Control*, *59*(11), 2873–2885.

Microsoft. n.d. Excel Performance - Improving Calculation Performance. https://docs.microsoft.com/en-us/office/vba/excel/concepts/excel-performance/excel-improving-calcuation-performance. [Accessed September 14, 2017].

Monteith, J. L. 1972. Solar radiation and productivity in tropical ecosystems. *Journal of Applied Ecology*, 9(3), 747–766.

Monteith, J. L. 1996. The Quest for Balance in Crop Modeling. *Agronomy Journal*, 88(5), 695. https://doi.org/10.2134/agronj1996.00021962008800050003x

Monteith, J. L., C. K. Ong, and J. E. Corlett. 1991. Microclimatic interactions in agroforestry systems. *Forest Ecology and Management*, 45(1–4), 31–44. https://doi.org/10.1016/0378-1127(91)90204-9

Monteith, J. L. and M. Unsworth. 2013. *Principles of environmental physics: plants, animals, and the atmosphere*. Academic Press.

Murrell, T. S. 2013. On-Farm Trials and Statistics. In *Proceedings Western Nutrient Management Conference*. Vol. 10. http://www.ipni.net/ipniweb/conference/wnmc.nsf/0/211da715452ce0aa85257bf8004df04d/$FILE/WNMC2013%20Murrell%20pg131.pdf.

Nash, J. C. 2010. Parameter Scaling Optimization and Related Nonlinear Modelling Computations in R. https://www.r-project.org/conferences/useR-2010/tutorials/Nash.pdf. [Accessed June 7, 2018].

Nielsen, R. L. B. 2010. *A Practical Guide to On-Farm Research.* https://www.agry.purdue.edu/ext/corn/news/timeless/OnFarmResearch.pdf. [Accessed February 18, 2017].

Oreskes, N., K. Shrader-Frechette, and K. Belitz. 1994. Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263(5147), 641–646.

Palma, J. H. N. 2015. CliPick: Project Database of Pan-European Climate Data for Default Model Use. Milestone Report 26 (6.1) for EU FP7 Research Project: AGFORWARD 613520. 10 October 2015. 22 pp.

Palma, J. H. N. 2017. CliPick–Climate Change Web Picker. A Tool Bridging Daily Climate Needs in Process Based Modelling in Forestry and Agriculture. *Forest Systems* 26 (1): 14. doi: 10.5424/fs/2017261-10251

Palma, J. H. N., A. R. Graves, J. Crous-Duran, M. Upson, J. A. Paulo, T. S. Oliveira, S. Silvestre Garcia de Jalón, P. J. Burgess,. 2016. Yield-SAFE Model Improvements. Milestone Report 29 (6.4) for EU FP7 Research Project: AGFORWARD 613520. (5 July 2016). 30 pp.

Palma, J. H. N., T. Oliveira, J. Crous-Duran, A. R. Graves, S. Garcia de Jalon, M. Upson, M. Giannitsopoulos, P. J. Burgess, J. A. Paulo, M. Tomé, N. Ferreiro-Domínguéz, M. R. Mosquera-Losada, P. Gonzalez-Hernández, S. Kay, J. Mirk, M. Kanzler, J. Smith, G. Moreno, A. Pantera, D. Mantovani, A. Rosati, B. Luske, J. Hermansen, J. 2017. Deliverable 6.17 (6.2): Modelled agroforestry outputs at field and farm scale to support biophysical and environmental assessments AGFORWARD project. 18 October 2017. 162 pp.

Palma, J. H. N., R. M. Cardoso, P. M. M. Soares, T. S. Oliveira, and M. Tomé. 2018. Using High-Resolution Simulated Climate Projections in Forest Process-Based Modelling. *Agricultural and Forest Meteorology* 263 (December): 100–106. https://doi.org/10.1016/j.agrformet.2018.08.008.

Parillo, P., and L. Ljung. 2003. Initialization of physical parameter estimates. Division of Automatic Control, Department of Electrical Engineering, Linkopings Universitet. Retrieved from http://www.diva-portal.org/smash/record.jsf?pid=diva2:316776

Passioura, J. B. 1973. Sense and nonsense in crop simulation. *J. Aust. Inst. Agric. Sci*, 39(3), 181–183.

Pearson, R. K. 2006. Nonlinear Empirical Modeling Techniques. *Computers & Chemical Engineering*, Papers form Chemical Process Control VII CPC VII Seventh international conference in the Series, 30 (10–12): 1514–28. doi:10.1016/j.compchemeng.2006.05.028.

Pease, C. M., and J. J. Bull. 1992. Is Science Logical? *Bioscience*; Oxford, 42(4), 293.

Pérez-Cruzado, C., F. Muñoz-Sáez, F. Basurco, G. Riesco, and R. Rodríguez-Soalleiro. 2011. Combining empirical models and the process-based model 3-PG to predict *Eucalyptus nitens* plantations growth in Spain. *Forest Ecology and Management*, 262(6), 1067–1077. https://doi.org/10.1016/j.foreco.2011.05.045

Pongnumkul, S., P. Chaovalit, and N. Surasvadi. 2015. Applications of Smartphone-Based Sensors in Agriculture: A Systematic Review of Research. *Journal of Sensors* 2015 (July): e195308. doi:10.1155/2015/195308.

Popper, K. 2005. *The Logic of Scientific Discovery*. Routledge.

Poushter, J. 2016. Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies. *Pew Research Center*. http://www.diapoimansi.gr/PDF/pew_research%201.pdf.

Pretzsch, H. 2009. Forest Dynamics, Growth, and Yield. In Forest Dynamics, Growth and Yield (pp. 1–39). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-540-88307-4_1

Pretzsch, H., D. I. Forrester, and T. Rötzer. 2015. Representation of species mixing in forest growth models. A review and perspective. *Ecological Modelling*, 313, 276–292. https://doi.org/10.1016/j.ecolmodel.2015.06.044

Priesack, E., and S. Gayler. 2009. Agricultural Crop Models: Concepts of Resource Acquisition and Assimilate Partitioning. In *Progress in Botany* (pp. 195–222). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-68421-3_9

Pugnaire, F. I., P. Haase, and J. Puigdefabregas. 1996. Facilitation between higher plant species in a semiarid environment. *Ecology*, *77*(5), 1420–1426.

Raue, A., C. Kreutz, T. Maiwald, U. Klingmüller, and J. Timmer. 2011. Addressing Parameter Identifiability by Model-Based Experimentation. *IET Systems Biology* 5 (2): 120–30. https://doi.org/10.1049/iet-syb.2010.0061.

Reichert, P., and M. Omlin. 1997. On the Usefulness of Overparameterized Ecological Models. *Ecological Modelling* 95 (2): 289–99. doi:10.1016/S0304-3800(96)00043-9.

Saltelli, A., and P. Annoni. 2010. How to Avoid a Perfunctory Sensitivity Analysis. *Environmental Modelling & Software* 25 (12): 1508–1517.

Schilders, W. 2008. Introduction to model order reduction. In *Model order reduction: Theory, research aspects and applications* (pp. 3–32). Springer.

Schillinger, W. F. 2011. Practical Lessons for Successful Long-Term Cropping Systems Experiments. *Renewable Agriculture and Food Systems* 26 (1): 1.

Schmid, S., E. Thürig, E. Kaufmann, H. Lischke, and H. Bugmann. 2006. Effect of forest management on future carbon pools and fluxes: A model comparison. *Forest Ecology and Management*, 237(1), 65–82.

Schoeneberger, M. M., G. Bentrup, and T. Patel-Weynand. 2017. *Agroforestry: Enhancing Resiliency in U.S. Agricultural Landscapes under Changing Conditions*;

General Technical Report WO-96; U.S. Department of Agriculture, Forest Service: Washington, DC, USA. doi.org/10.2737/WO-GTR-96.

Sedmák, R., and L. Scheer. 2015. Properties and prediction accuracy of a sigmoid function of time-determinate growth. *Iforest—Biogeosciences and Forestry*, e1–e7. https://doi.org/10.3832/ifor1243-007

Shvets, V., and B. Zeide. 1996. Investigating parameters of growth equations. *Canadian Journal of Forest Research*, 26(11), 1980–1990. https://doi.org/10.1139/x26-224

Sinclair, T. R., and N. G. Seligman. 1996. Crop Modeling: From Infancy to Maturity. *Agronomy Journal*, 88(5), 698. https://doi.org/10.2134/agronj1996.00021962008800050004x

Sivakumar, B. 2008. Dominant Processes Concept, Model Simplification and Classification Framework in Catchment Hydrology. *Stochastic Environmental Research and Risk Assessment* 22 (6): 737–48. doi:10.1007/s00477-007-0183-5.

Sjöberg, J., Q. Zhang, L. Ljung, A. Benveniste, B. Delyon, P.-Y. Glorennec, H. Hjalmarsson, and A. Juditsky. 1995. Nonlinear Black-Box Modeling in System Identification: A Unified Overview. *Automatica*, Trends in System Identification, 31 (12): 1691–1724. doi:10.1016/0005-1098(95)00120-8.

Slock, D.T.M. 1993. On the Convergence Behavior of the LMS and the Normalized LMS Algorithms. *IEEE Transactions on Signal Processing* 41 (9): 2811–25. https://doi.org/10.1109/78.236504.

Sooby, J. 2001. On-Farm Research Guide. Organic Farming Research Foundation. http://www.ofrf.org/sites/ofrf.org/files/docs/pdf/on-farm_research_guide_rvsd.pdf.

Stage, A. R. 2003. How Forest Models Are Connected to Reality: Evaluation Criteria for Their Use in Decision Support. *Canadian Journal of Forest Research* 33 (3): 410–21. doi:10.1139/x02-203.

Thacker, W. C. 1989. The Role of the Hessian Matrix in Fitting Models to Measurements. *Journal of Geophysical Research* 94 (C5): 6177. https://doi.org/10.1029/JC094iC05p06177.

Toensmeier, E. 2016. *The Carbon Farming Solution: A Global Toolkit of Perennial Crops and Regenerative Agriculture Practices for Climate Change Mitigation and Food Security*. Chelsea Green Publishing: Hartford, VT, USA.

USGCRP. 2017. *Climate Science Special Report: Fourth National Climate Assessment, Volume I* [Wuebbles, D. J., D. W. Fahey, K. A. Hibbard, D. J. Dokken, B. C. Stewart, and T. K. Maycock (eds.)]. U.S. Global Change Research Program, Washington, DC, USA, 470 pp, doi: 10.7930/J0J964J6.

Valentine, H. T., and A. Mäkelä. 2005. Bridging process-based and empirical approaches to modeling tree growth. *Tree Physiology*, 25(7), 769–779.

van der Ploeg, R. R., W. Böhm, and M. B. Kirkham, and others. 1999. On the Origin of the Theory of Mineral Nutrition of Plants and the Law of the Minimum. https://dl.sciencesocieties.org/publications/sssaj/abstracts/63/5/1055.

van der Werf, W., K. Keesman, P. J. Burgess, A. Graves, D. Pilbeam, L. D. Incoll, ... and C. Dupraz. 2007. Yield-SAFE: A parameter-sparse, process-based dynamic model for predicting resource capture, growth, and production in agroforestry systems. *Ecological Engineering*, 29(4), 419–433. https://doi.org/10.1016/j.ecoleng.2006.09.017

van Ittersum, M. K., and M. Donatelli. 2003. Modelling Cropping Systems—Highlights of the Symposium and Preface to the Special Issues. *European Journal of Agronomy* 18 (3–4): 187–97. doi:10.1016/S1161-0301(02)00095-3.

van Ittersum, M.K., and R. Rabbinge. 1997. Concepts in Production Ecology for Analysis and Quantification of Agricultural Input-Output Combinations. *Field Crops Research* 52 (3): 197–208. doi:10.1016/S0378-4290(97)00037-3.

van Meijgaard, E., L. H. Van Ulft, G. Lenderink, S. R. De Roode, E. L. Wipfler, R. Boers, and R. M. A. van Timmermans. 2012. *Refinement and Application of a Regional Atmospheric Model for Climate Scenario Calculations of Western Europe*. KVR 054/12. KVR.

van Noordwijk, M., B. Lusiana, N. Khasanah, and R. Mulia. 2011. WaNuLCAS Version 4.0, Background on a Model of Water Nutrient and Light Capture in Agroforestry Systems. World Agroforestry Centre. http://worldagroforestry.org/downloads/WaNuLCAS/WaNuLCAS4.0.pdf.

Vanclay, J. K. 1994. Modelling forest growth and yield: applications to mixed tropical forests. School of Environmental Science and Management Papers, 537. Wallingford, UK: CAB International.

Vastaranta, M., E. G. Latorre, V. Luoma, N. Saarinen, M. Holopainen, and J. Hyyppä. 2015. Evaluation of a Smartphone App for Forest Sample Plot Measurements. *Forests* 6 (4): 1179–1194.

Verchot, L. V., M. V. Noordwijk, S. Kandji, T. Tomich, C. Ong, A. lain Albrecht, J. Mackensen, C. Bantilan, K.V. Anupama, and C. Palm. 2007. Climate Change: Linking Adaptation and Mitigation through Agroforestry. *Mitig. Adapt. Strategies Glob. Chang. 12*, 901–918.

Vilela, M., S. Vinga, M. A. G. M. Maia, E. O. Voit, and J. S. Almeida. 2009. Identification of neutral biochemical network models from time series data. *BMC Systems Biology*, *3*(1), 47. https://doi.org/10.1186/1752-0509-3-47

Villalobos, F. J., and E. Fereres (eds.). 2016. *Principles of Agronomy for Sustainable Agriculture*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-46116-8

Villaverde, A. F., A. Barreiro, and A. Papachristodoulou. 2016. Structural Identifiability of Dynamic Systems Biology Models. *PLOS Computational Biology*, 12(10), e1005153. https://doi.org/10.1371/journal.pcbi.1005153

Vries, F. P. de. 1989. *Simulation of ecophysiological processes of growth in several annual crops* (Vol. 29). Int. Rice Res. Inst.

Walter, E. 2012. Identifiability, and Beyond. In *System Identification, Environmental Modelling, and Control System Design* (pp. 49–68). Springer.

Wang, D., and M. Murphy. 2004. Estimating optimal transformations for multiple regression using the ACE algorithm. *Journal of Data Science*, *2*(4), 329–346.

Waterworth, R. M., G. P. Richards, C. L. Brack, and D. M. W. Evans. 2007. A generalised hybrid process-empirical model for predicting plantation forest growth. *Forest Ecology and Management*, 238(1), 231–243. https://doi.org/10.1016/j.foreco.2006.10.014

Weiskittel, A. R. 2014. Forest Growth and Yield Models for Intensively Managed Plantations. In J. G. Borges, L. Diaz-Balteiro, M. E. McDill, & L. C. E. Rodriguez (eds.), The Management of Industrial Forest Plantations (pp. 61–90). Springer Netherlands. https://doi.org/10.1007/978-94-017-8899-1_3

Weiskittel, A. R., D. W. Hann, J. A. Kershaw Jr, and J. K. Vanclay. 2011. *Forest Growth and Yield Modeling*. John Wiley & Sons.

Whisler, F. D., B. Acock, D. N. Baker, R. E. Fye, H. F. Hodges, J. R. Lambert, ... V. R. Reddy. 1986. Crop Simulation Models in Agronomic Systems. *Advances in Agronomy*, 40, 141–208. https://doi.org/10.1016/S0065-2113(08)60282-5

Whittaker, G., R. Confesor, M. Di Luzio, and J. G. Arnold. 2010. Detection of Overparameterization and Overfitting in an Automatic Calibration of SWAT. *Transactions of the ASABE* 53 (5): 1487–1499.

Widrow B., and S. D. Stearns. 1985. *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall.

Yao, K. Z., B. M. Shaw, B. Kou, K. B. McAuley, and D. W. Bacon. 2003. Modeling Ethylene/Butene Copolymerization with Multi-Site Catalysts: Parameter

Estimability and Experimental Design. *Polymer Reaction Engineering* 11 (3): 563–588.

Yin, X., and P. S. Struik, P. C. 2010. Modelling the crop: from system dynamics to systems biology. *Journal of Experimental Botany*, 61(8), 2171–2183. https://doi.org/10.1093/jxb/erp375

Young, P. 1983. The validity and credibility of models for badly defined systems. In *Uncertainty and Forecasting of Water Quality* (pp. 69–98). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-82054-0_2

Young, P. C. 2011. *Recursive Estimation and Time-Series Analysis*. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-21981-8

Young, P. C. 2012. Data-Based Mechanistic Modelling: Natural Philosophy Revisited? In: Wang, L., & Garnier, H. *System Identification, Environmental Modelling, and Control System Design* (pp. 321–340). Springer.

Young, P. C. 2013. Hypothetico-inductive data-based mechanistic modeling of hydrological systems. *Water Resources Research*, 49(2), 915–935. https://doi.org/10.1002/wrcr.20068

Young, P. C., M. Ratto. 2009. A unified approach to environmental systems modeling. *Stochastic Environmental Research and Risk Assessment*, 23(7), 1037–1057. https://doi.org/10.1007/s00477-008-0271-1

Zadeh, L. A. 1962. From Circuit Theory to System Theory. *Proceedings of the IRE*, 50(5), 856–865. https://doi.org/10.1109/JRPROC.1962.288302

Zarco-Tejada, P. J., Luciano M., E. Fereres, and F. J. Villalobos. 2016. New Tools and Methods in Agronomy. In *Principles of Agronomy for Sustainable Agriculture*, pp. 503-514. Springer: Cham.

Zeide, B. 1993. Analysis of growth equations. *Forest Science*, 39(3), 594–616.

Ziehn, T., and A. S. Tomlin. 2009. GUI–HDMR–A Software Tool for Global Sensitivity Analysis of Complex Models. *Environmental Modelling & Software* 24 (7): 775–785.